

文章编号:1007-2780(2018)04-0347-10

基于神经网络集成模型的宫颈细胞病理 计算机辅助诊断方法

廖欣^{1,2}, 郑欣³, 邹娟^{1,2}, 冯敏^{1,2}, 孙亮^{1,2}, 杨帆^{1*}

(1. 四川大学 华西第二医院 病理科, 四川 成都 610041;

2. 四川大学 出生缺陷与相关妇科疾病教育部重点实验室, 四川 成都 610041;

3. 电子科技大学 计算机科学与工程学院, 四川 成都 611731)

摘要: 针对宫颈细胞病理图像自动筛查问题, 本文提出一种基于人工智能技术的计算机辅助诊断方法。该方法通过对宫颈细胞病理图像采用自适应双阈值法进行初步检测, 再采用改进 Chan-Vase 模型进行精确分割, 提取出细胞(粘连簇团)中的不同区域。然后, 结合病理诊断专家规则, 构建相应的正交特征集。在此基础上, 使用神经网络集成模型进行正常、疑似病变二分类识别, 完成计算机辅助诊断。实验表明, 本文方法能够有效完成宫颈病理细胞(粘连簇团)的分类识别, 具有较高的正确率(84%)与较低的误判率(2.1%)。满足了在保证判断正确率的条件下, 尽量降低将疑似病变样本误判为正常样本的实际病理诊断要求。

关键词: 宫颈; 细胞病理; 筛查; 神经网络集成; 计算机辅助诊断

中图分类号: TP394.1; TH691.9 **文献标识码:** A **doi:** 10.3788/YJYXS20183304.0347

Computer-aided diagnosis of cervical cytopathology based on neural network ensemble model

LIAO Xin^{1,2}, ZHENG Xin³, ZOU Juan^{1,2}, FENG Min^{1,2}, SUN Liang^{1,2}, YANG Fan^{1*}(1. *Department of Pathology, West China Second University Hospital,*
*Chengdu 610041, China;*2. *Key Laboratory of Birth Defects and Related Disease of Women and Children,*
*Ministry of Education, Sichuan University, Chengdu 610041, China;*3. *College of Computer Science and Engineering, University of Electronic*
Science and Technology of China, Chengdu 611731, China)

Abstract: Aiming to the automatic screening of cervical cytopathological images, an artificial intelligence based automatic diagnosis-assisted method was proposed. First of all, adaptive dual threshold method was used to detect the cervical cytopathological images initially. Secondly, improved Chan-

收稿日期: 2017-12-16; 修订日期: 2018-02-06.

基金项目: 四川省重点实验室开放基金(No.2017LF3008); 广东省应用型研发重大专项基金(No.2015BD10131002)

Supported by Program of Key Laboratory Open Fund in Sichuan Province(No. 2017LF3008); The Important Special Fund for Applied R & D in Guangdong Province(No.2015BD10131002)

* 通信联系人, E-mail: xinty927@163.com

Vase model was used to precisely extract different areas of adhesive cell cluster. After that, the related feature set was built according to the diagnostic rules of pathological experts. At last, neural network ensemble was applied to normal or suspected lesions two-classification recognition. The result of the experiment showed that cervical cell lesions could be effectively distinguished according to classification with this method, which had high accuracy (84%) and low rate of misjudgment (2.1%), meeting the practical requirement of pathological diagnosis, which is reducing the miscalculating of the suspected lesions to normal ones, meanwhile assuring the diagnostic accuracy.

Key words: cervix; cytopathology; screening test; neural network ensemble; computer-aided diagnosis

1 引言

早期筛查是对宫颈癌进行预防控制的关键途径^[1-2]。我国宫颈癌早期普查工作量十分繁重,然而病理医生数量却严重不足。因此,利用人工智能技术进行宫颈细胞病理辅助诊断在宫颈癌前病变诊断中具有重要意义^[3-4]。基于Thirna提出的细胞图像处理及相关量化方法的框架,美国研发了相关系统:PapNet系统^[4]借助自动显微镜筛选出128幅含有相对异常细胞的图像区域,供病理医生着重阅览诊断;AutoPaP系统^[5]、FocalPoint系统^[6]设定一个包含多种细胞参数的综合评价指标;ThinPrep系统^[7]能排除病理图像中25%的正常区域,但筛查准确率均有限。高细见、全宇、赵英红等^[8-10]研究者分别针对宫颈细胞图像分割方法进行了研究,由于宫颈细胞涂片粘连严重,导致实际分割效果有限。李光研究了宫颈癌细胞核特征及诊断意义^[11]。国内针对宫颈细胞病理自动识别的研究大多集中于算法本身,如宫颈细胞图像分割,细胞特征提取,以及对单个细胞进行识别分类的理论研究阶段,缺乏针对宫颈细胞病理辅助诊断系统实现方法的相关研究。

本文提出了基于神经网络集成模型的宫颈细胞病理计算机辅助诊断方法。首先利用初检测精分割的方式提取细胞病理图像中的兴趣区域。然后,结合病理诊断专家知识以及图像分析技术,针对兴趣区域构造识别特征,并通过PCA算法进行特征选择、降维。最后,采用神经网络集成模型完成识别分类。

2 宫颈细胞病理计算机辅助诊断流程

针对宫颈细胞病理图像自动识别方法进行研究,结合细胞病理诊断知识,运用人工智能、图像分析技术完成宫颈细胞的分类识别,其流程如图1所示。

宫颈细胞病理图像识别包括以下环节:

- (1)提取宫颈病理细胞(粘连簇团)区域;
- (2)细胞(粘连簇团)区域有效识别特征的构造、选择;
- (3)基于神经网络集成的智能识别模型。

此外,本文研究的核心是对宫颈细胞病理图像进行自动识别,针对细胞粘连簇团区域中各个细胞的精确分离超出本文的研究范围,但将是我们今后的研究重点。

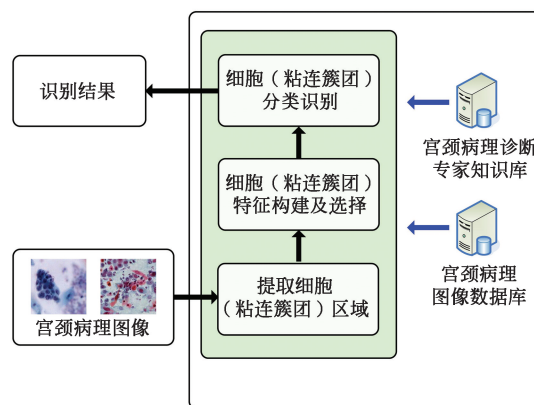


图1 宫颈细胞病理图像自动识别流程

Fig. 1 Flow chart of automatic recognition of cervical cytopathological image

3 病理图像中的细胞检测及特征提取

3.1 宫颈病理细胞(粘连簇团)区域提取

针对宫颈细胞病理图像的特点,我们采用先初步检测,再精确分割的方法来提取宫颈病理细胞(粘连簇团)区域,其技术路线如图2所示,具体描述如下:

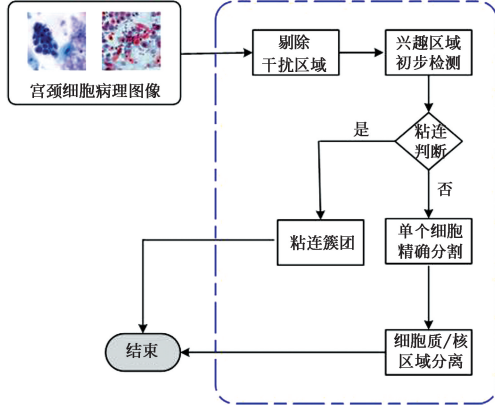


图2 宫颈病理细胞(粘连簇团)提取流程

Fig.2 Flow chart of extract cervical cytopathological image

步骤1、初检测采用自适应双阈值算法^[12],设定 T_1 、 T_2 初始值。

$$\begin{aligned} T_1 &= (2U_1 + U_2) / 3, \\ T_2 &= (U_1 + 2U_2) / 3, \end{aligned} \quad (1)$$

其中:第一次迭代时, U_1 取图像的最小灰度 a , U_2 取图像的最大灰度 b ;根据式(2)将图像分割成 R_1 、 R_2 、 R_3 三个区域。

$$\begin{cases} a \leq x_i \leq T_1, x_i \in R_1 \\ T_1 \leq x_i \leq T_2, x_i \in R_2, \\ T_2 \leq x_i \leq b, x_i \in R_3 \end{cases} \quad (2)$$

计算各区域均值 U_1 、 U_2 、 U_3 ,并按 $T_1 = (U_1 + U_2) / 2$, $T_2 = (U_2 + U_3) / 2$ 更新阈值 T_1 , T_2 。重复以上步骤,直到 T_1 、 T_2 不再变化。采用最终阈值 T_1 、 T_2 对原图像进行分割,完成初步检测。

步骤2、针对初步检测结果提取连通区域,构建候选细胞(粘连簇团)区域,并根据区域面积判断候选区域中包含单个细胞还是粘连细胞簇团。

步骤3、采用改进 Chan-Vase 模型^[13]精确分割候选区域。调整 λ 可分离出单个细胞区域中的细胞核、细胞体、背景区域,以及粘连细胞簇团

区域中的细胞簇团、细胞核区域。采用基于区域内一致性(方差),区域间差异性(灰度均值之差的平方)的改进 C-V 模型的能量泛函如下:

$$\begin{aligned} E(C) &= (1-\lambda) \times E_1(C) - \lambda \times E_2(C) + \\ &\quad \mu \times E_3(C) + \nu \times E_4(C) = \\ &\quad (1-\lambda) \times \left[\iint_{\Omega} |\mu_0 - c_1|^2 H(\varphi) dx dy + \right. \\ &\quad \left. \iint_{\Omega} |\mu_0 - c_2|^2 [1 - H(\varphi)] dx dy \right] - \\ &\quad \frac{\lambda}{2} \times |c_1 - c_2|^2 + \mu \int_{\Omega} \delta(\varphi) |\nabla(\varphi)| dx dy + \\ &\quad \nu \int_{\Omega} H(\varphi) dx dy, \end{aligned} \quad (3)$$

采用两个独立水平集,通过设置不同加权系数 λ ,分别逼近细胞核区域、核外部区域的边界 B_1 ,细胞体区域、背景区域的边界 B_2 ,用以分离背景、胞核、胞体。加权系数 λ 为 0 时,为基于区域内一致性的 Chan-Vase 模型; λ 为 1 时,为基于区域间差异性的 Chan-Vase 模型,其中:

$$\begin{aligned} E_1(C) &= \iint_{\Omega} |\mu_0 - c_1|^2 H(\varphi) dx dy + \\ &\quad \iint_{\Omega} |\mu_0 - c_2|^2 [1 - H(\varphi)] dx dy, \end{aligned} \quad (4)$$

$$E_2(C) = \frac{\lambda}{2} \times |c_1 - c_2|^2, \quad (5)$$

$$\begin{aligned} E_3(C) &= \text{length}(\varphi = 0) = \\ &\quad \int_{\Omega} \delta(\varphi) |\nabla(\varphi)| dx dy, \end{aligned} \quad (6)$$

$$E_4(C) = \text{Area}(\varphi \geq 0) = \int_{\Omega} H(\varphi) dx dy, \quad (7)$$

$E_1(C)$ 为基于区域内一致性的能量函数, $E_2(C)$ 为基于区域间差异性的能量函数, $E_3(C)$ 为闭合轮廓曲线长度约束, $E_4(C)$ 为区域面积约束。Heaviside 函数 H , 一维 Dirac 函数 R 。水平集函数迭代完成后,提取零水平集($t=0$)得到最终轮廓曲线。

步骤4、候选细胞(粘连簇团)区域中可能存在一个以上的连通区域,改进 Chan-Vase 模型迭代停止后,常得到多于一条的候选细胞体轮廓、细胞核轮廓。需要根据面积阈值从候选细胞体轮廓区域中判断主细胞体、主细胞核,或主粘连簇团区域。

3.2 宫颈病理细胞特征提取及选择

包括宫颈细胞特征的构建、提取,以及特征降

维及归一化,技术路线如图 3 所示。结合细胞病理诊断知识,以及图像分析技术,我们选择的特征类型以及具体特征(5 类 24 种特征)如表 1 所示,其定义及计算方式参考文献[14-17]。

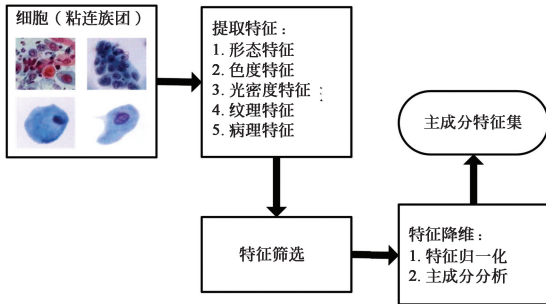


图 3 构建有效识别特征流程

Fig.3 Flow chart of building effective recognition feature

表 1 宫颈病理细胞的候选特征

Tab.1 Candidate feature of cervical cell lesions

特征类型	特征参数	
形态特征	面积	周长
	面积/周长比	等效圆直径
	等价半径	圆形成度
	矩形度	伸长度
色度特征	RGB 分量均值	RGB 分量方差
	RGB 分量色度变化系数	
光密度特征	平均光密度	光密度方差
	光密度变化系数	光密度偏度
	光密度峰态	光密度对比度
	光密度对比度	光密度对比度
纹理特征 (Tamura)	粗糙度	对比度
	线性度	规整度
病理特征	特征距离均匀度	核变形态
	核内部染色度	

分类识别所用的特征并非越多越好,引入无效特征反而会增大分类误差。因此,首先进行特征归一化,消除数量级差别,使不同特征具有可比性。采用主成分分析法构建主成分特征集,特征矩阵 $\mathbf{X}$:

$$\mathbf{X} = (\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_p) = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix}, \quad (8)$$

设均值 $E(X_k) = \mu_k$, 方差 $Var(X_k) = \sigma_{kk}$, 获得

标准化特征变量 X_k^*

$$X_k^* = \frac{X_k - \mu_k}{\sqrt{\sigma_{kk}}}, \quad k = 1 \sim p, \quad (9)$$

构建协方差矩阵,并按照公式(10)选取前 m 个 ($m < p$) 最大特征值 λ_i 对应的特征向量构成主成分特征集^[18]。

$$\begin{cases} \sum_{i=1}^{m-1} \lambda_i < 80\% \sim 85\% \\ \sum_{i=1}^p \lambda_i \\ \sum_{i=1}^m \lambda_i \geq 80\% \sim 85\% \\ \sum_{i=1}^p \lambda_i \end{cases}. \quad (10)$$

3.3 基于神经网络集成的智能识别模型

采用神经网络集成构建宫颈细胞病理图像智能识别模型,以实现细胞(粘连簇团)区域的正常、疑似病变两分类。同时,结合病理诊断规则,设计相关识别策略,以提高智能识别模型的实际诊断价值,流程如图 4 所示,步骤如下。

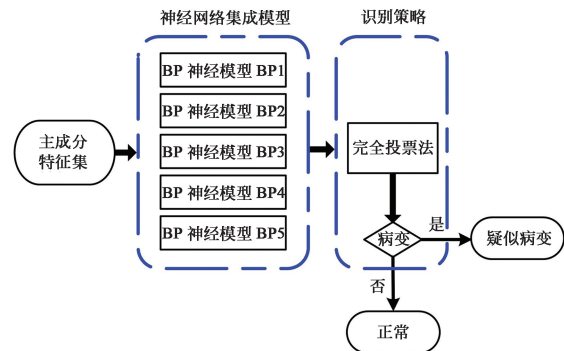


图 4 神经网络集成模型识别流程

Fig.4 Flow chart of neural network ensemble model

步骤 1、单个神经网络选择 BP 神经网络模型,其传递函数采用对数型 Sigmoid 函数,训练函数采用 trainrp 函数。个体神经网络输入层神经元数目设置为主成分特征集的特征维数 L ,采用 1 个隐含层(神经元数目 $2L$),输出层采用 1 个神经元(二值输出)。

步骤 2、采用 Bagging 集成技术^[19] (Bootstrap Sampling),生成 5 个独立的 BP 神经网络。Bagging 算法如图 5 所示,从大小为 N 的原始训练集 D 中,通过对样本集随机、有放回采样 N 次,产生一个与原始训练集大小相同的训练样本子集,训

练同构基分类器。各训练轮间相互独立,增加网络集成的差异度,能提高系统泛化能力,适合并行训练多个基础分类器。

步骤3、设计智能识别模型的识别策略时,需要满足在保证整体识别率的前提下,降低将病变细胞误判为正常细胞的概率。因此,采用完全投票法策略判断正常细胞,即仅当集成模型中的五个(考虑模型识别性能及复杂程度的经验优选值)BP神经网络同时判断为正常时,才识别其为正常类别。

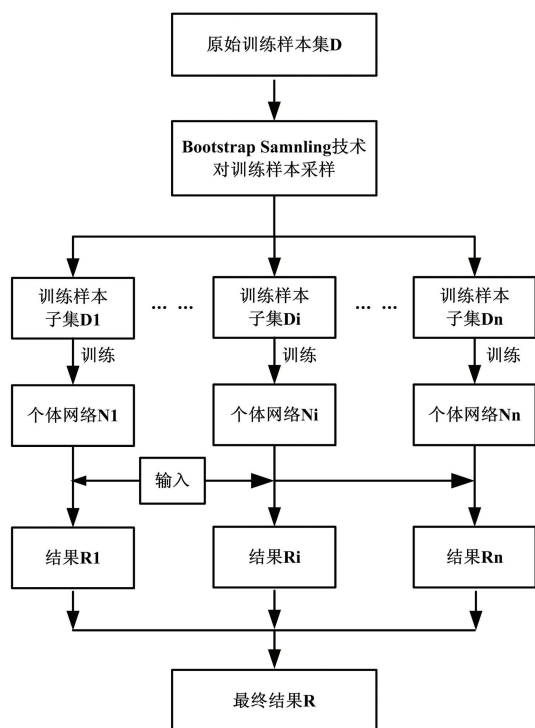


图5 Bagging集成技术生成个体神经网络

Fig.5 Schematic diagram of individual neural network with Bagging integrated technology

4 实验及分析

为了验证本文方法的有效性,我们利用自建宫颈细胞病理图像库进行了相关实验。

4.1 自建宫颈细胞病理图像库

一定数量的宫颈细胞病理图像,是进行自动识别方法研究的基础。因此,我们建立了宫颈细胞病理图像数据库,共计140幅20倍物镜下的宫颈鳞状细胞的镜下图像,包含了正常、低度病变、高度病变、非典型细胞(粘连簇团)等类型,图6所示是其中4幅典型图像。库中图像来源包括本院

病理确诊的宫颈细胞病理LCT涂片扫描图像,以及Bethesda分级标准中的部分标准图像。

针对自建宫颈细胞病理图像库中的图像,我们完成了以下工作:(1)选择病理图像时,排除了病理涂片中的干扰,包括宫颈腺细胞、细胞重叠、常见炎症、出血区域等;(2)由三甲医院病理科主诊医生对图像中的细胞(粘连簇团)完成第三方标注,即进行正常、疑似病变二分类;其中,疑似病变类别包含了Bethesda分级标准中的高度鳞状病变、低度鳞状病变、非典型鳞状细胞等类型。

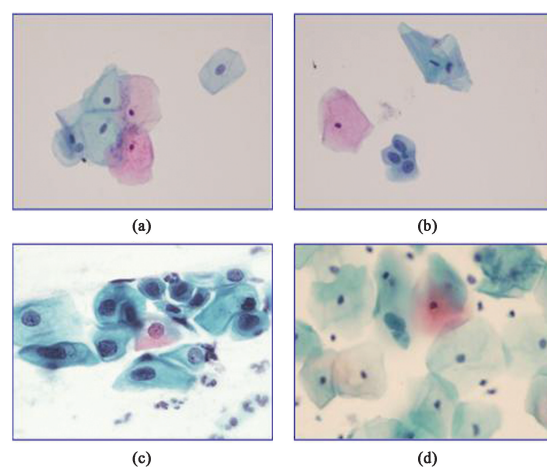


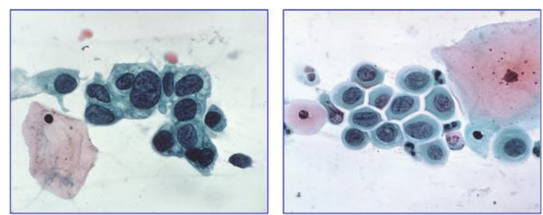
图6 神经网络集成模型识别流程

Fig.6 Flow chart of neural network ensemble model

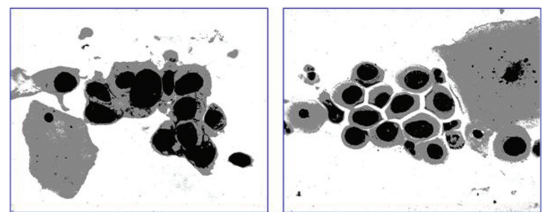
4.2 宫颈细胞病理检测实验及分析

图7展示了自适应双阈值算法的检测过程,其中图7(a)是原始宫颈细胞病理图像,图7(b)是自适应双阈值算法检测结果,图7(c)是采用形态学开运算消除毛刺后的结果。其中,细胞核区域表示为黑色,细胞质区域表示为灰色,背景区域表示为白色。

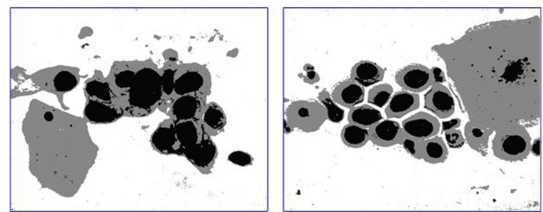
采用改进Chan-Vase模型从初步检测出的候选区域中进行精确分割,设置细胞核初始轮廓曲线C1、细胞体初始轮廓曲线C2的半径分别为候选区域短边长度的1/6、1/3,且均以候选区域中心为圆心的圆周线。改进Chan-Vase模型完成迭代后,首先剔除不封闭的轮廓线,保留封闭轮廓线,使用阈值保留细胞体轮廓线、细胞核轮廓线,并采用面积最大的轮廓作为主轮廓。图8、图9分别展示了单个细胞、粘连细胞簇团的精确提取过程。



(a)宫颈细胞病理图像
(a)Image of cervical cytopathology



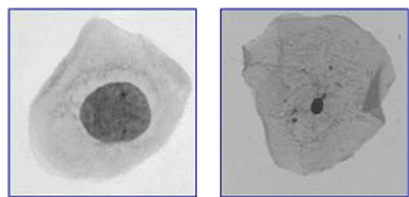
(b)全局初步检测结果
(b)Result of preliminary detection



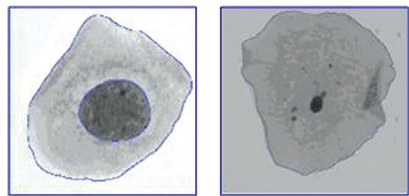
(c)形态学处理结果
(c)Result of result of morphotogical processing

图 7 宫颈细胞病理图像初步检测

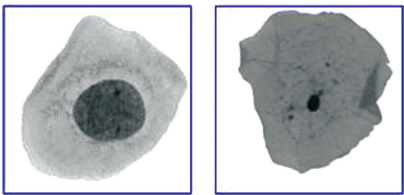
Fig.7 Initial detection of cervical cytopathology image



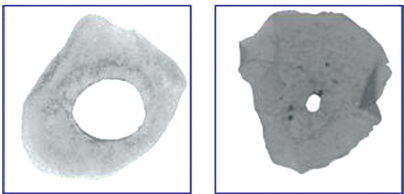
(a)单个细胞
(a)Single cell



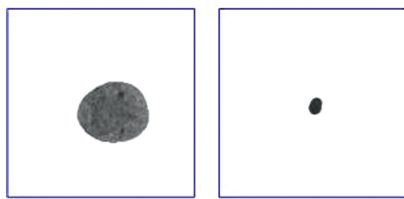
(b)细胞主轮廓
(b)Main contour of cell



(c)细胞体
(b)Cell body



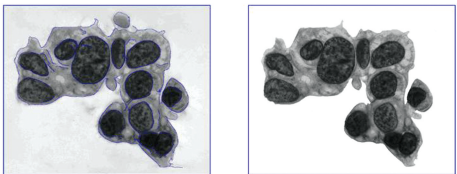
(d)细胞质
(d)Cytoplasm



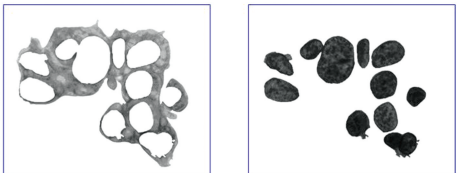
(e)细胞核
(e)Cell nucleus

图 8 单个宫颈病理细胞精确提取

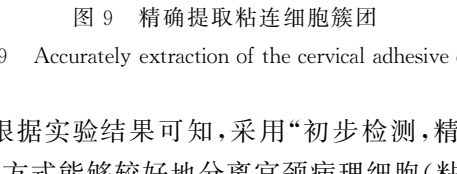
Fig.8 Accurately extraction of a single cervical cell



(a)迭代结束轮廓
(a)Iterated end outline



(a)细胞质提取结果
(a)Result of cytoplasm extraction



(d)细胞核提取结果
(d)Result of cell nucleus extraction

图 9 精确提取粘连细胞簇团

Fig.9 Accurately extraction of the cervical adhesive cells

根据实验结果可知,采用“初步检测,精确提取”的方式能够较好地分离宫颈病理细胞(粘连簇

团)区域,达到了预期目标,是实现自动识别的基础。从自建宫颈细胞病理图像库中共得到单个细胞496个,其中正常细胞268个,疑似病变细胞228个;粘连细胞簇团区域112个,其中正常区域47个,疑似病变区域65个,如表2所示。

表2 图像库中的宫颈病理细胞(粘连簇团)计数

Tab.2 Amount of cervical lesion cells (adhesive clusters) in image library

	正常	疑似病变	总计
单个细胞	268	228	496
粘连细胞簇团	47	65	112
总计	315	293	608

4.3 宫颈细胞病理自动识别实验及分析

保证整体识别正确率的前提下,需要降低将

病变细胞(粘连簇团)误判为正常细胞(粘连簇团)的概率。因此,神经网络集成模型采用完全投票法,当5个BP神经网络全部判断细胞(粘连簇团)为正常类别,集成模型才将其判断为正常类别,否则判断为疑似病变类别。

采用3.2节方法,针对表1列出的24种特征进行PCA变换,并根据公式(10)选择变换后,对应特征值最大的8个特征向量组成主成分特征集,用于后续BP神经网络模型集成识别。实验中单个BP模型的输入神经元数量设置为8,采用一个隐层,隐层中BP神经元数量设置为16,输出层采用2个BP神经元,分别对应正常、疑似病变两种类别。根据上述实验设置,分别针对单个细胞样本集、粘连细胞簇团样本集,以及全体样本集(同时包含单个细胞、粘连细胞簇团)进行特征提取、选择,然后采用神经网络集成模型进行分类识别,实验结果如表3所示。

表3 神经网络集成模型在不同样本集上的识别率

Tab.3 Recognition rate of network ensemble model in different sample sets

		单个细胞 样本集	粘连细胞簇团 样本集	全体 样本集
BP1	识别正确率	82.6%	85.7%	83.2%
	正常样本误识别率	20.1%	12.8%	19.0%
	疑似样本误识别率	14.0%	15.4%	14.3%
BP2	识别正确率	84.9%	87.5%	85.4%
	正常样本误识别率	11.9%	21.3%	13.3%
	疑似样本误识别率	18.9%	6.2%	16.0%
BP3	识别正确率	86.3%	88.4%	86.7%
	正常样本误识别率	16.0%	10.6%	16.2%
	疑似样本误识别率	11.0%	12.3%	11.3%
BP4	识别正确率	87.7%	86.6%	87.8%
	正常样本误识别率	10.1%	17.0%	11.1%
	疑似样本误识别率	14.0%	15.4%	14.3%
BP5	识别正确率	89.7%	92.0%	90.1%
	正常样本误识别率	7.8%	8.5%	7.9%
	疑似样本误识别率	13.2%	7.7%	11.9%
神经网络 集成模型	识别正确率	83.7%	88.4%	84.0%
	正常样本误识别率	26.1%	12.5%	13.8%
	疑似样本误识别率	4.8%	1.8%	2.1%

表 3 中列出了神经网络集成模型,以及其中的 5 个 BP 神经网络模型针对单个细胞样本集、粘连细胞簇团样本集、全体样本集的相关识别率,包括样本识别正确率、正常样本误识别率、疑似样本误识别率。通过对比、分析表 3 中的数据,得到以下结论:

(1)神经网络集成模型针对单个细胞样本集的识别正确率为 83.7%,相比单个 BP 模型的最佳性能 89.7%有一定下降,略低于 5 个 BP 神经网络模型的平均性能 86.2%。相较于单个 BP 神经网络的将疑似样本误判为正常样本概率的最优性能(最小值 11.0%),神经网络集成模型的此项误判率仅为 4.8%,其性能得到明显提高。

(2)神经网络集成模型针对粘连细胞簇团样本集的识别正确率为 88.4%,相比单个 BP 模型的最佳性能 92.0%略微下降,接近于 5 个 BP 神经网络模型的平均性能 88.0%。相较于单个 BP 神经网络的将疑似样本误判为正常样本概率的最优性能(最小值 6.2%),神经网络集成模型的此项误判率仅为 1.8%,其性能得到明显提高。

(3)神经网络集成模型针对全体样本集的识别正确率 84.0%,相比单个 BP 神经网络模型的最佳性能 90.1%有一定下降,接近于 5 个 BP 模型的平均性能 86.6%。相较于单个 BP 神经网络的将疑似样本误判为正常样本概率的最优性能(最小值 11.3%),神经网络集成模型的此项误判率仅为 2.1%,其性能得到明显提高。

基于上述分析可知,神经网络集成模型的识别正确率相对单个 BP 神经网络集成模型的最佳性能略有下降,基本接近于 5 个 BP 神经网络模型的平均识别率。但集成模型将疑似样本判断为正常样本的误识别率明显下降,即使相对于单个 BP 神经网络模型的最佳性能进行对比。由于宫

颈细胞病理图像的背景复杂,以及细胞(粘连簇团)区域存在特征多样性的特点,对其进行精确提取的难度较大,识别准确率较低^[20]。我们采用神经网络模型集成,利用单个 BP 神经网络模型针对不同特征的细胞(粘连簇团)区域具有较佳识别能力的特点,再通过完全投票法对识别结果进行集成。实验表明,该方法针对不同样本集(单个宫颈病理细胞样本集、粘连细胞簇团样本集、全体样本集)均有稳定表现,能够取得较好的分类识别性能。同时,在保证样本整体识别准确率的提前下,显著降低了将疑似样本误判为正常样本的可能性。

5 结 论

所提宫颈细胞病理计算机辅助诊断方法,通过自适应双阈值算法对细胞(粘连簇团)区域进行初检测,再采用改进 Chan-Vase 模型进行精确分割,克服了病理图像背景复杂,细胞个体存在差异的困难。然后针对宫颈细胞病理图像特点,结合细胞病理诊断知识,构建了用于后续分类识别的特征集。采用主成分分析法构建主成分特征集,消除冗余特征。在此基础上,采用神经网络集成模型进行自动识别,在保证识别准确率的前提下,显著降低了将疑似样本误判为正常样本的概率。在自建宫颈细胞病理图像库中的实验结果表明,本文方法的识别准确率达到 84%,将疑似病变样本误判为正常样本的错误率仅为 2.1%。

本文在一定程度上解决了宫颈细胞病理图像的自动识别问题,但针对细胞粘连簇团中单个细胞的精确分离还有待深入研究,导致其实用性受到一定限制。因此,我们后续将重点研究针对粘连细胞簇团中细胞的精确分离方法。

参 考 文 献:

- [1] ASHTARIAN H, MIRZABEIGI E, MAHMOODI E, *et al.* Knowledge about cervical cancer and Pap smear and the factors influencing the pap test screening among women [J]. *International Journal of Community Based Nursing and Midwifery*, 2017, 5(2): 188-195.
- [2] KIM J J, BURGER E A, SY S, *et al.* Optimal cervical cancer screening in women vaccinated against human papillomavirus [J]. *Journal of the National Cancer Institute*, 2017, 109(2): djw216.
- [3] 胡佐鸿,赵春,包骥,等.全片数字化图像在诊断细胞病理学中的应用[J].*中华病理学杂志*,2017,46(8):581-585.
HU Z H, ZHAO C, BAO J, *et al.* Application of whole slide imaging in diagnostic cytology [J]. *Chinese Journal of Pathology*, 2017, 46(8): 581-585. (in Chinese)

- [4] JENNY J, ISENEGGER I, BOON M E, *et al.* Consistency of a double PAPNET scan of cervical smears [J]. *Acta Cytologica*, 1997, 41(1): 82-87.
- [5] BIBBO M, HAWTHORNE C. Performance of the AutoPap primary screening system at Jefferson University Hospital [J]. *Acta Cytologica*, 1999, 43(1): 27-29.
- [6] KARDOST F. The FocalPoint system: FocalPoint slide profiler and FocalPoint GS [J]. *Cancer Cytopathology*, 2004, 102(6): 334-339.
- [7] HUTCHINSON M L, ISENSTEIN L M, GOODMAN A, *et al.* Homogeneous sampling accounts for the increased diagnostic accuracy using the ThinPrep™ processor [J]. *American Journal of Clinical Pathology*, 1994, 101(2): 215-219.
- [8] 高细见,曾立波,吴琼水,等.一种基于显微多光谱宫颈细胞图像自动分割方法[J].数据采与处理,2004,19(4): 441-445.
- GAO X J, ZENG L B, WU Q S, *et al.* Automatic segmentation method for cervical cell based on multispectral microgram [J]. *Journal of Data Acquisition & Processing*, 2004, 19(4): 441-445. (in Chinese)
- [9] 全宇,王忠庆,何苗.基于交叉熵的神经网络在病理图像分析中的应用[J].中国医科大学学报,2009,38(6):446-448.
- QUAN Y, WANG Z Q, HE M. Application of neural network based on cross-entropy method in pathological image analysis [J]. *Journal of China Medical University*, 2009, 38(6): 446-448. (in Chinese)
- [10] 赵英红,周中顺,孙存杰,等.一种新的用于宫颈癌粘连细胞图像分割的分水岭算法[J].军事医学,2014,38(12): 972-975.
- ZHAO Y H, ZHOU Z S, SUN C J, *et al.* A new modified watershed algorithm for image segmentation in cervical cancer cell adhesion[J]. *Military Medical Sciences*, 2014, 38(12): 972-975. (in Chinese)
- [11] 李光,张海峰,王军梅,等.宫颈鳞状细胞癌细胞核的形态定量分析[J].山西医科大学学报,2005,36(4):429-431.
- LI G, ZHANG H F, WANG J M, *et al.* Nuclear morphometrical analysis of cervix cancer [J]. *Journal of Shanxi Medical University*, 2005, 36(4): 429-431. (in Chinese)
- [12] LIANG G M, TANG C J, LIU D H, *et al.* A multi-level and adaptive dual threshold segmentation algorithm base on evaluation [J]. *Acta Electronica Sinica*, 2009, 37(4): 750-752, 763.
- [13] 王相海,王金玲,方玲玲.灰度差能量函数引导的图像分割自适应 C-V 模型[J].模式识别与人工智能,2015,28(3): 214-222.
- WANG X H, WANG J L, FANG L L. An adaptive C-V image segmentation model guided by gray difference energy function[J]. *Pattern Recognition and Artificial Intelligence*, 2015, 28(3): 214-222. (in Chinese)
- [14] 赵爱昱,王宏力,杨小冈,等.纹理粗糙度在红外图像显著性检测中的应用[J].光学精密工程,2016,24(1): 220-228.
- ZHAO A G, WANG H L, YANG X G, *et al.* Application of texture coarseness in saliency detection of infrared image[J]. *Optics and Precision Engineering*, 2016, 24(1): 220-228. (in Chinese)
- [15] 纪慧中,贾大宇,董思清,等.基于图像特征和光流场的非刚性图像配准[J].光学精密工程,2017,25(9): 2469-2482.
- JI H Z, JIA D Y, DONG E Q, *et al.* Non-rigid registrations based on image characteristics and optical flows[J]. *Optics and Precision Engineering*, 2017, 25(9): 2469-2482. (in Chinese)
- [16] 于萍,步宏,王华,等.免疫组化结果的图像分析与人工计数方法的对比研究[J].生物医学工程学杂志,2003,20(2):288-290.
- YU P, BU H, WANG H, *et al.* Comparative study on image analysis and manual counting of immunohistochemistry[J]. *Journal of Biomedical Engineering*, 2003, 20(2): 288-290. (in Chinese)
- [17] 王宇,邹文辉,杨晓敏,等.基于计算机视觉的 X 射线图像异物分类研究[J].液晶与显示,2017,32(4):287-293.
- WANG Y, ZOU W H, YANG X M, *et al.* X-ray image illegal object classification based on computer vision [J]. *Chinese Journal of Liquid Crystals and Displays*, 2017, 32(4): 287-293. (in Chinese)
- [18] ROWEIS S. EM algorithms for PCA and SPCA [C]//*Proceedings of the 1997 Conference on Advances in Neural Information Processing Systems*. Denver, Colorado, USA: ACM, 1997: 626-632.
- [19] ALFARO M E, ZOLLER S, LUTZONI F. Bayes or bootstrap? A simulation study comparing the performance of

- Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence [J]. *Molecular Biology and Evolution*, 2003, 20(2): 255-266.
- [20] ZHANG L, LU L, NOGUES I, *et al.* DeepPap: Deep convolutional networks for cervical cell classification [J]. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21(6): 1633-1643.

作者简介:廖欣(1981—),女,四川成都人,硕士,病理主治医师;研究方向:妇产科临床病理诊断;人工智能在疾病诊断与病理分析中的应用等;E-mail:xinty927@163.com

郑欣(1981—),男,四川绵阳人,博士后;研究方向:健康大数据;人工智能在大健康及医疗领域的应用等;E-mail:1820745320@qq.com

杨帆(1975—),男,四川南溪人,博士,病理副主任医师;研究方向:妇科肿瘤及诊断;E-mail:fanyang@scu.edu.cn