

局部二维结构描述的 HEp-2 染色模式分类

张东波^{1,2}, 颜 霜¹, 张 莹^{1,2}, 秦 海¹, 王俊超¹

(1. 湘潭大学信息工程学院, 湖南湘潭 411105; 2. 机器人视觉感知与控制技术国家工程实验室, 湖南湘潭 411105)

摘 要: HEp-2 染色模式分类主要用于免疫疾病诊断, 但已有方法受荧光成像环境, 细胞图像自身的视觉特性的影响, 分类准确率较低. 本文提出一种新的适合于 HEp-2 染色模式分类的特征提取方法. 在构建了不同尺度下的高斯平滑图像序列后, 利用 shape index 实现图像二维结构的直观描述, 进而通过多灰度阈值图像结构的空分分解, 使其同时具备对微观二维图像结构和空间信息的描述能力. 该方法在 ICPR 和 SNP HEp-2 数据集的两折交叉细胞级测试中, 分别获得 89.83% 和 87.49% 的准确率, 在 ICPR 的 28 折交叉细胞级和图像级测试分别达到 60.5% 和 70.56% 的准确率, 明显优于 LBP、CLBP 等方法, 和 CoALBP 特征相当.

关键词: 间接免疫荧光图像; 染色模式; 空间分解

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112 (2016)12-2817-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2016.12.001

HEp-2 Staining Pattern Classification by Local 2D Structure Description

ZHANG Dong-bo^{1,2}, YAN Shuang¹, ZHANG Yin^{1,2}, QIN Hai¹, WANG Jun-chao¹

(1. College of Information Engineering, Xiangtan University, Xiangtan, Hunan 411105, China;

2. Robot Visual Perception & Control Technology National Engineering Laboratory, Xiangtan, Hunan 411105, China)

Abstract: HEp-2 staining pattern classification is used for immune disease diagnosis. Due to the influence of imaging environment and the visual characteristics of cell images, the recognition accuracy of reported methods are still unsatisfied. A novel feature suitable for HEp-2 staining pattern classification is proposed. After constructing serial smoothing images with Gaussian scale parameter, the description ability of shape index is used to describe second-order image structure, moreover, the space structure can be presented by multi-threshold image segmentation. The proposed approach was tested in ICPR and SNPHEp-2 datasets by 2-fold cross validation in cell level, and achieved 89.83% and 87.49% accuracy for the two datasets respectively. Moreover, in 28 fold cross validation test of ICPR, the method achieved 60.5% and 70.56% accuracy in cell level and image level respectively. Experimental results show that the proposed method is superior to other popular texture descriptors, such as LBP and CLBP, and approximate to the performance of CoALBP feature.

Key words: indirect immune fluorescence (IIF) image; staining pattern; space decomposition

1 引言

间接免疫荧光(IIF)图像分析是现代医学对自身免疫疾病进行诊断的有效手段,不同的IIF荧光染色模式与不同的免疫疾病密切相关.荧光染色模式的准确分类目前仍然是一个非常困难的问题,原因在于受荧光成像环境影响,不同染色模式细胞图像外观差异变化大,而其中一些染色模式又具有非常相似的视觉特性.

HEp-2 细胞染色模式分类采用的特征主要有形状/形态学特征、纹理特征二类.例如 G. V. Ponomarev^[1]提

取 ROI 染色细胞区域的数量,大小,形状等形态学特征, R. Stoklasa^[2]通过不同尺度形态学开运算描述细胞图像中的峰值区域分布特性,同时借助地形学理论,对图像灰度分布中呈现出来的类似山、谷形状的表面信息进行统计.作为一种表面形状描述子,该特征计算简单,具有优越的分类性能.此外,由于荧光响应模式通常呈现出规律性的分布和空间关系,因此纹理特征也是一种非常有效的对 HEp-2 染色模式分类的特征,但全局纹理特征,例如灰度共生矩阵 GLCM 和离散余弦变换 DCT^[3],不能很好的描述图像中的微观结构之间的空间

关系,因此分类性能欠佳.近年来,局部纹理特征得到广泛应用,例如 Haralick 特征^[2],LBP 特征^[4-7],SIFT 特征^[8],其中 LBP 特征具有计算成本低,容易实现,鲁棒性较好等优点,且由于擅长描述图像微观结构,LBP 及其改进的特征在 HEp-2 染色模式分类中得到有效应用,例如 R. Nosaka^[7]考虑了成对 LBPs 之间空间关系提出的 CoALBP 特征,为了实现旋转不变性,提高鲁棒性,提出了 RIC-LBP 特征. I. Theodorakopoulos^[8]结合 SIFT 与 GoC-LBPs 提取特征,其中,SIFT 特征具有旋转不变性,对光照变化,噪声,平移以及缩放具有鲁棒性,而 GoC-LBPs 特征考虑了各模式之间信息共现性,两种特征的结合很好的描述了细胞图像的灰度和纹理信息. S. Di. Cataldo^[9]则提出了形态学与纹理特征结合的特征描述方法,其形态学特征考虑连通区域圆度(即面积/周长²)、数量、比值以及面积等等,纹理特征则考虑了全局纹理特征(GLCM,ZERN,EOH,RIGF)和局部纹理特征(LBPriu2,CLBP,CoLBP, RIC-LBP).组合特征虽然一定程度能够提升分类性能,但同时加大了计算量,算法复杂,而性能提升非常有限.

由 J. J. Koenderink 提出的 shape index^[10]是一种能够对图像中的二维结构进行直观描述的特征,受此启发,本文基于 shape index 并结合尺度空间理论和基于灰度排序的空间分解提出了一种新的特征提取方法.该特征具备对图像微观二维结构和空间信息的描述能力,实验结果验证了新特征的良好分类能力.

2 本文方法

为避免图像尺度对统计特征的影响,细胞图像均归一成 64×64 的尺寸,同时为避免光照影响通过式(1)还做了灰度归一化变换.

$$I' = \frac{I - \text{mean}(I)}{\text{std}(I)} \quad (1)$$

其中, $\text{mean}(I)$ 和 $\text{std}(I)$ 分别是细胞图像的均值和标准差, I 是原图像, I' 预处理后图像.

2.1 基于 shape index 的特征提取

Shape index 是由 Koenderink 和 van Doorn^[10]提出的一种图像几何测量方法.由于 HEp-2 图像中的二维结构有尺寸差异,因此有必要在高斯尺度空间框架内计算 shape index. 图像的高斯尺度空间,

$$L(\mathbf{x}; \sigma) = (G * I')(\mathbf{x}; \sigma), \mathbf{x} = (x, y) \quad (2)$$

其中, $*$ 表示卷积运算, G 是高斯函数

$$G(\mathbf{x}; \sigma) = \frac{1}{(\sigma \sqrt{2\pi})^2} \exp\left(-\frac{\mathbf{x}^T \mathbf{x}}{2\sigma^2}\right), \sigma > 0 \quad (3)$$

其中 σ 为高斯核函数的尺度参数.通过规范化尺度空间导数^[11]可以计算图像中的差分结构

$$L_{x^m y^n}(\mathbf{x}; \sigma) = \sigma^{n+m} \frac{\partial^{n+m}}{\partial x^n \partial y^m} (G * I')(\mathbf{x}; \sigma)$$

$$= \sigma^{n+m} \left(\left(\frac{\partial^{n+m}}{\partial x^n \partial y^m} G \right) * I' \right) (\mathbf{x}; \sigma) \quad (4)$$

其中 n 和 m 分别代表的沿 x 轴和 y 轴方向的差分阶次,为了描述简便,尺度 σ 下的差分值 $L_{xy}^{nm}(\mathbf{x}; \sigma)$ 记作 L_{xy}^{nm} .

通过计算二阶导数 L_{x^2}, L_{y^2}, L_{xy} 我们可以构造出如下的 Hessian 矩阵:

$$\nabla^2 L(\mathbf{x}; \sigma) = \begin{bmatrix} L_{x^2} & L_{xy} \\ L_{xy} & L_{y^2} \end{bmatrix} \quad (5)$$

Hessian 矩阵是二阶对称矩阵,其特征值 k_1 和 k_2 可以用来衡量尺度 σ 下的二阶曲率,由下式计算,

$$k_1(\mathbf{x}; \sigma) = \frac{1}{2} (L_{x^2} + L_{y^2} - \sqrt{(L_{x^2} - L_{y^2})^2 + 4L_{xy}^2}) \quad (6)$$

$$k_2(\mathbf{x}; \sigma) = \frac{1}{2} (L_{x^2} + L_{y^2} + \sqrt{(L_{x^2} - L_{y^2})^2 + 4L_{xy}^2}) \quad (7)$$

k_1 和 k_2 称之为主曲率,分别描述的是曲率在极大值和极小值方向的值,值得注意的是, k_1 和 k_2 的值本身和极值曲率方向无关,因此具有旋转不变性.当 $k_1 = k_2 = 0$ 时,此时的像素点区域平坦,无明显信息结构.为了得到描述图像二维结构具有旋转不变性的特征,定义如下 shape index 指标 $s \in [-\pi/2, \pi/2]$,

$$s(\mathbf{x}; \sigma) = \arctan\left(\frac{k_1 + k_2}{k_1 - k_2}\right) \quad (8)$$

指标 s 具有旋转,灰度以及平移不变性,能描述连续图像空间区间内所有的二阶形状.当 $s = 0$ 时,对应的是一个鞍状结构;当 $s \neq 0$ 时,则可以分别描述极大值、极小值、脊线和谷线等凹凸不一的曲面结构.

s 指标只能描述曲面结构的形状,而相同形状的结构在图像中的可观测性不一样,有的易于观测,有的和背景接近很难区分,因此二维结构的显著性需要另外用曲度值 c (curvedness) 进行区分

$$c(\mathbf{x}; \sigma) = \sqrt{k_1^2 + k_2^2} \quad (9)$$

曲度描述的所描绘的强度的强度, c 值小意味着此形状几乎是平坦的且不明显, c 值大则意味着此形状是显著的.

联合式(6)~(9)可以得到计算 shape index 的 s 和 c 的公式如下,

$$s(\mathbf{x}; \sigma) = \arctan\left(\frac{-L_{x^2} - L_{y^2}}{\sqrt{(L_{x^2} - L_{y^2})^2 + 4L_{xy}^2}}\right) \quad (10)$$

$$c(\mathbf{x}; \sigma) = \sqrt{L_{x^2}^2 + 2L_{xy}^2 + L_{y^2}^2} \quad (11)$$

测量值 s 和 c 可看作一类简明且有效的纹理描述方法,类似于 SIFT 特征构建的梯度方向直方图,我们可以将 s 看作梯度方向, c 看作梯度值,作为直方图统计时梯度方向的权重,借鉴这样一种处理策略,我们不难实现图像中和表面形状有关的纹理信息描述.

为构建多尺度描述下的 shape index,本文采用了如

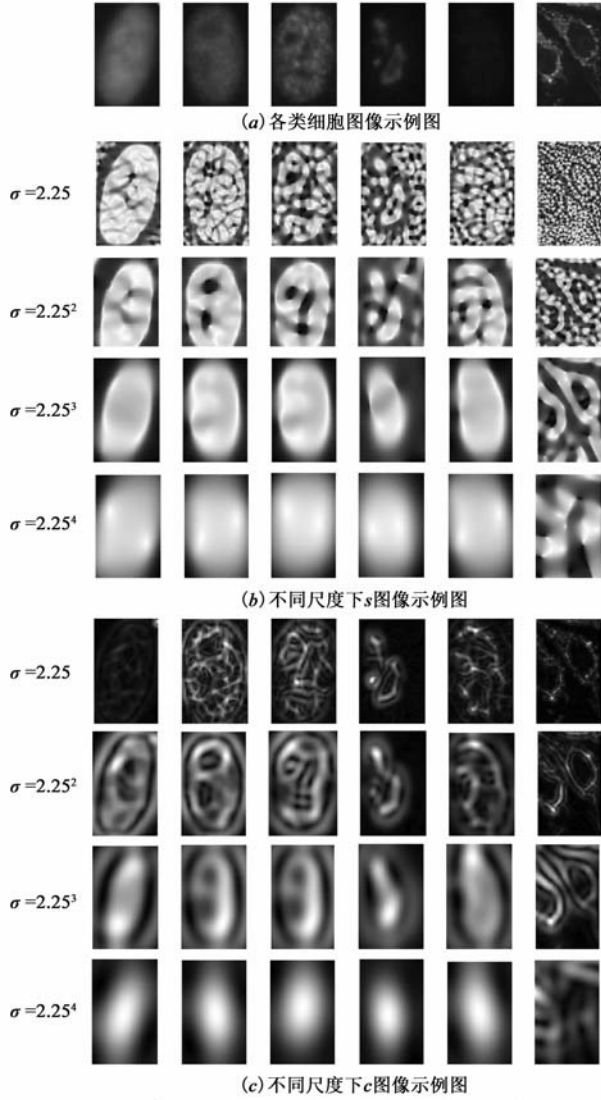


图1 不同尺度下各类细胞图像的主曲率方向和曲率示例图

下多尺度(n_σ)的选择原则

$$\sigma_i = \sigma_{\text{base}} \cdot \sigma_{\text{ratio}}^i, i = 1, \dots, n_\sigma \quad (12)$$

其中, σ_{base} 是 shape index 的最小基准尺度, σ_{ratio} 是相邻尺度 σ_i 和 σ_{i+1} 的比值, $\sigma_{\text{ratio}}^{i-1}$ 表示 σ_{ratio} 的 $i-1$ 次方. 图 1(b) 和图 1(c) 分别给出了不同模式类细胞图像的曲率方向 s 及曲率 c 在不同观测尺度时的示例图. 从图 1 可以看出, 曲率方向 s 示意图用于观测图像中的二维结构分布情况, 当尺度较小时, 更多反映的是图像细小的纹理和斑点区域, 随尺度增大, 整体轮廓信息得以凸显. 而曲度 c 示意图, 在尺度较小时, 大曲度信息主要出现在图像中细小结构的边缘, 随尺度增大, 大曲度信息会逐渐集中出现在细胞图像的内部, 由于不同染色模式的内部斑点分布和整体轮廓有差异, 因此随尺度变化, 不难观测到各种模式的 s 示意图和 c 示意图的变化规律呈现显著差异.

2.2 空间分解策略

shape index 作为一种局部特征, 为获得良好的分辨能力, 需要在特征描述时根据图像特点对图像空间做适当的分解以便刻画重要的空间信息. 为获得旋转不变性, 我们考虑了两种基于灰度信息的空间分解策略.

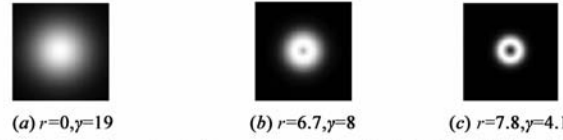
2.2.1 基于环状结构的空分分解

第一种策略为同心环状结构空分分解策略, 以点 x_0 为中心, 半径为 r 的环状区域 D 描述如下,

$$D(\mathbf{x}; \mathbf{x}_0, \gamma, r) = \exp\left(-\frac{(\|\mathbf{x}_0 - \mathbf{x}\|_2 - r)^2}{2\gamma^2}\right) \quad (13)$$

其中 γ 是高斯核函数宽度. 通过上述公式, 我们可以对到中心点距离为 r 的环状区域中各点按高斯函数衰减特性赋予一定的权重.

为了提高特征描述能力, 我们可以选择多组参数 n_s , 以便实现对空间的多重分解, 通过实验我们最终选择了三组参数, 即 $(r, \gamma) = (0, 19)$, $(6.7, 8)$ 和 $(7.8, 4.1)$ (如图 2 所示). 该组参数可以描述具有交叠的同心环状结构空间, 能较好的覆盖了整个细胞图像区域.

图2 选择三组 r 和 γ 参数, 即 $n_s=3$ 时的环状空分分解结构示意图

2.2.2 基于 Ostu 阈值分割的空分分解

首先根据 Ostu 方法得到一个参考阈值 TH4, 然后以 TH4 为中心, 分别在小于和大于 TH4 的灰度区间再等间隔选择 3 个阈值, 总共得到 7 个分割阈值, 通过 7 个阈值我们可以得到 7 个空分分解子图 ($n_s = 7$). 图 3 分别给出了 Homogeneous, Fine Speckled, Coarse Speckled, Nucleolar, Centromere 和 Cytoplasmatic 六类模式下的细胞示例的阈值分割结果. 显然, 对于不同的染色模式, 其 7 组阈值分割结果图具有明显不同的变化规律, 其中隐含了丰富的染色模式内部结构信息.

2.3 直方图构建

为了建立基于 shape index 指标的直方图, 在 $[-\pi/2, \pi/2]$ 区间均匀选取 n_b 个等级, 且对这 n_b 个区间分别取它们的中心值 b , 然后使用空分分解权重 D 和主曲率值 c 构建直方图. 对于第一种空分分解策略, 图像中各点依据式 (13) 选取, 而对于第二种空分分解策略, 图像中分割出来的前景区域各点的权重 D 都设为 1, 背景设为 0, 最后依据加权求和, 计算整幅图像在给定参数时的直方图分布, 计算公式如下,

$$B(\delta, b, \beta) = \frac{\sum_x D c \exp\left(-\frac{(b-s)^2}{2\beta^2}\right)}{\sum_x D} \quad (14)$$

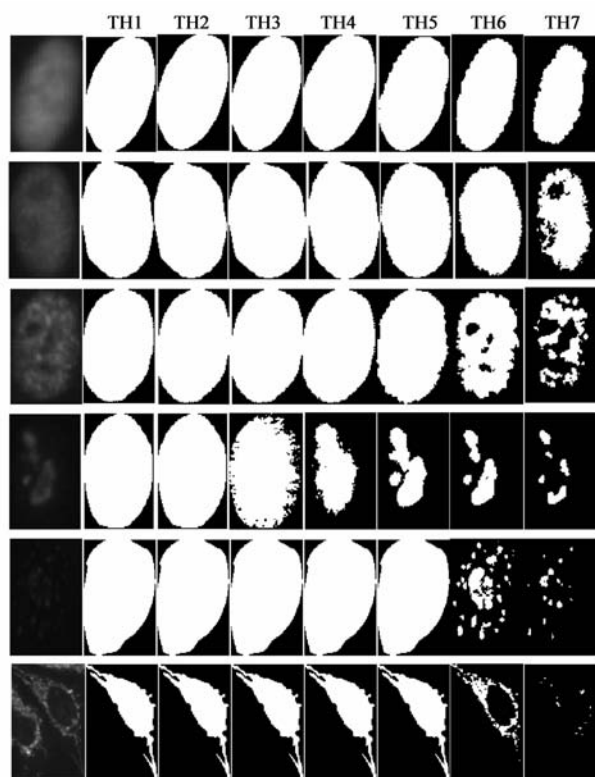


图3 各类染色模式细胞TH1~TH7掩膜图像

实验中, $n_b = 18$, 参数 b 是这 n_b 个区域中的中心值 (取值范围在 $[-17\pi/36, 17\pi/36]$). β 是高斯核函数宽度. 根据上述公式可以得到在不同尺度 σ 和不同空间结构 D 下的 shape index 直方图, 通过拼接获得整体直方图特征向量. 对于第一种和第二种空间分解策略, 我们分别最终可以得到 $n_s \cdot n_b \cdot n_s = 216$ 和 $n_s \cdot n_b \cdot n_s = 504$ 维的特征向量. 最后, 为了对抗对比度变化, 需要对特征向量做归一化, 根据实验测试, 本文采用的是 L_1 归一化方法.

3 实验结果与分析

为验证本文特征在 HEp-2 细胞模式分类中的效果, 我们在公开的数据集 ICPR2012^[12] 和 SNP 数据集^[13] 中进行测试, 采用的分类器是 LIBSVM 工具箱^[14], 评价准则用的是 [12] 推荐的两折交叉法和 28 折交叉校验.

LBP 作为一种擅于对图像微观结构进行描述的特征, 其计算原理简单, 具有旋转和光照不变性的显著特点, 因此是一种被广泛采用的纹理特征描述形式, 已有研究人员将其成功用于 HEp-2 细胞染色模式分类问题^[7], 并取得了良好的分类性能. 我们在实验中选择了旋转不变统一模式 LBPriu2 和完整局部二值模式 CLBP 两种典型的 LBP 特征和本文特征进行了对比. 为了找到合适的邻域参数提取 LBP 特征, 根据邻域半径 R 和邻域点数 P , 我们选取了三组参数 $(P, R) = (8, 1)$ 、 $(8,$

$2)$ 、 $(16, 2)$ 分别对 ICPR 和 SNP 数据集进行了测试. 当 (P, R) 取 $(16, 2)$ 这一组参数时, 不管是在 ICPR 还是 SNP 数据集, 分类性能都要优于其它两组参数, 此外, 由于 CLBP 特征不仅描述了邻域中像素点之间的灰度大小关系, 同时保留了其原有灰度信息, 因此分类性能优于 LBPriu2 特征, 以 (P, R) 取 $(16, 2)$ 为例, 在两个数据集中 CLBP-S-M 特征分类准确率分别比 LBPriu2 特征提升了 10.69% 和 15.17%.

此外, 为确定采用环状空间分解策略时, 相关参数 r 和 γ 的取值, 我们分别按 $n_s = 1, 2, 3$ 三种情况取了若干组参数, 并在 ICPR 数据集中作了测试. 在所有组测试中, 当 $n_s = 3$, 即 r 和 γ 分别取 $(0, 19)$, $(6.7, 8.0)$ 和 $(7.8, 4.1)$ 时, 获得了最好的分类准确率是 88.26%. 因此在后续实验中, 我们都是按此组参数来提取 shape index + donut 特征.

另外基准尺度参数 δ_{base} 和高斯核函数宽度参数 β 对于最终的 shape index + Ostu 特征分类性能也有重要影响, 对于 ICPR 数据集, 当 $\delta_{base} = 1.8$, $\beta = 0.4$ 时有最好分类性能 89.83%, 而对于 SNP 数据集, 当 $\delta_{base} = 2.0$, $\beta = 0.22$ 时有最好分类性能 87.49%.

表 1 给出了采用 LBPriu2 特征, CLBP-S-M 特征以及基于 shape index 采用不同空间分解策略的 shape index + donut 特征和 shape index + Ostu 特征在 ICPR 和 SNP 数据集中测得的两折交叉校验准确率, 其中 shape index 特征和 CLBP-S-M 分类性能相当, 但是加入空间分解策略后的 shape index + donut 特征和 shape index + Ostu 特征明显要优于单纯的 shape index 特征. shape index + Ostu 特征分类准确率在两个数据集中都要优于其它特征, 其中采用第二种空间分解策略之所以高于第一种策略的环状空间分解, 是因为环状空间分解人为弱化了相邻像素之间的灰度相似关系, 而基于 Ostu 阈值的空间分解, 较好地保留了像素点之间的灰度相似关系, 能够更好地描述图像中的二维结构, 因此其获得比环状空间分解更好的分类准确率.

表 1 HEp-2 细胞染色模式在各种特征下的分类准确率 (%)

数据集	LBPriu2	CLBP-S-M	shape index	shape index + donut	shape index + Ostu
ICPRContest	75.67	86.46	86.46	88.26	89.83
SNPHEp-2	69.99	83.98	83.88	85.26	87.49

ICPR 数据集采自 28 张切片图像, 每一幅图像的细胞染色模式基本属于同一类, 由于不同切片图像具有独立性, 因此 28 折交叉校验也是在 ICPR 数据集的测试中常用的评价准则, 通常我们使用其中 27 幅图像中的细胞做训练集, 剩下的一幅图像中的细胞做测试集, 依序处理做 28 次测试, 并取其平均值作为最终测试结果.

表 2 展示了使用 28 折交叉获得的各类图像细胞分类结果,其中标明了用于做检测的 28 幅图像的染色细胞的实际类别和预测类别,每类的正确分类和错误分类的数量和百分比.表 3 展示了在做图像级测试时的混淆矩阵,在图像级测试时,以该图像中细胞预测类别占多数的类作为该图的预测类别.从表 3 可以发现,在 Homogeneous 和 Fine Speckled 这两类中有很大的误分率,其中 Homogeneous 类有 40% 误分为 Fine Speckled 类, Fine Speckled 类有 75% 误分为 Homogeneous 类,具体表现在图像 ID 21, ID 22 中的 Homogeneous 被错分为 Fine

Speckled 的概率分别是 57.28% ,54.62% ,而正确分类的概率分别是 13.11% ,35.29% ;在图像 ID 9, ID 15, ID 23 中, Fine Speckled 被错分为 Homogeneous 的概率分别是 50% ,46.03% ,64.71% ,而正确分类的概率分别是 19.57% ,17.46% ,35.29% .可以看出这两类的错分率竟然到达了正确分类率的一半以上,原因在于这两类的图片非常相似,因此容易出现误分.而 Centromere 与 Cytoplasmatic 和其它模式比具有较明显的结构差异,易于区分,因此具有较高的分类准确率.

表 2 ICPR 数据集 28 折交叉校验的准确率(%)

IMAGE ID	实际类别	Hom 数量		C. Sp 数量		Nucl 数量		Cent 数量		F. Sp 数量		Cyt 数量	
01	Hom	59	96.72	0	0	0	0	0	0	2	3.28	0	0
02	F. Sp	4	8.33	11	22.9	0	0	1	2.08	32	66.67	0	0
03	Cent	0	0	4	4.49	0	0	78	87.64	0	0	7	7.87
04	Nucl	0	0	0	0	35	53.03	14	21.2	17	25.8	0	0
05	Hom	30	65.82	0	0	0	0	1	2.13	16	34.04	0	0
06	C. Sp	1	1.47	59	86.76	0	0	0	0	0	0	8	11.76
07	Cent	0	0	0	0	5	8.93	51	91.07	0	0	0	0
08	Nucl	2	3.57	0	0	8	14.29	13	23.21	0	0	33	58.93
09	F. Sp	23	50	14	30.43	0	0	0	0	9	19.57	0	0
10	C. Sp	0	0	20	60.60	0	0	2	6.06	11	33.33	0	0
11	C. Sp	1	2.44	21	51.22	0	0	0	0	19	46.34	0	0
12	C. Sp	1	2.04	44	89.8	0	0	1	2.04	0	0	3	6.12
13	Cent	0	0	3	6.52	0	0	43	93.47	0	0	0	0
14	Cent	7	11.11	7	11.11	3	4.76	5	7.94	41	65.08	0	0
15	F. Sp	29	46.03	5	7.94	7	11.11	10	15.87	11	17.46	1	1.59
16	Cent	1	2.63	0	0	3	7.89	34	89.47	0	0	0	0
17	C. Sp	2	10.53	1	5.26	0	0	0	0	16	84.21	0	0
18	Hom	30	73.17	4	9.76	0	0	0	0	7	17.07	1	2.44
19	Cent	0	0	0	0	5	7.69	60	92.31	0	0	0	0
20	Nucl	5	10.87	0	0	37	80.43	4	8.70	0	0	0	0
21	Hom	8	13.11	0	0	3	4.92	15	24.6	35	57.38	0	0
22	Hom	42	35.29	5	4.20	1	0.84	0	0	65	54.62	6	5.04
23	F. Sp	33	64.71	2	3.92	1	1.96	0	0	15	29.41	0	0
24	Nucl	7	9.59	6	8.22	57	78.08	1	1.37	2	2.74	0	0
25	Cyt	1	4.17	1	4.17	4	16.67	4	16.67	5	20.83	9	37.5
26	Cyt	1	2.94	5	14.7	0	0	1	2.94	0	0	27	79.41
27	Cyt	0	0	1	2.63	0	0	0	0	0	0	37	97.37
28	Cyt	0	0	0	0	0	0	1	7.69	1	7.69	11	84.62

表 3 图像级检测时的混淆矩阵(%)

	图像级分类					
	Hom	C. Sp	Nucl	Cent	F. Sp	Cyt
Hom	60	0	0	0	40	0
C. Sp	0	80	0	0	20	0
Nucl	0	0	75	0	0	25
Cent	0	0	0	83.33	16.67	0
F. Sp	75	0	0	0	25	0
Cyt	0	0	0	0	0	100

表 4 其他方法与本文方法对比(%)

方法	LBP	LBP ^{ri}	LBP ^{riu2}	CLBP-S-M	CoALBP	shape index + Ostu
细胞级	49.48	47.07	51.75	56.90	59.03	60.50
图像级	57.14	57.14	57.14	64.28	71.40	70.56

表 4 中给出了常规 LBP 和若干改进型的最新的 LBP 特征描述方法和本文方法的比较,结果是在采用 28 折交叉校验时,分别在细胞级和图像级基础上测得的染色模式分类准确率,其中 CoALBP 和本文方法要明

显优于其它特征,CoALBP 在细胞级和图像级测试中,分别达到 59.03% 和 71.4% 的准确率,而本文的 shape index + Ostu 特征同样达到了 60.50% 和 70.56% 的准确率,值得说明的是,CoALBP 特征在 2012 年的 ICPR 竞赛中获得了第一名的成绩,而本文提出的特征达到了和它相当的性能,这也充分验证了本文方法的有效性和分类性能.表 5 总结了本文最终采用的 shape index + Ostu 特征在两个数据集上采用不同准则的测试结果.

表 5 本文最终算法测试结果总结 (%)

数据集	参数	两折交叉		28 折交叉	
		细胞级	图像级	细胞级	图像级
ICPRContest	$\delta_{\text{base}} = 1.8, \beta = 0.4$	89.83	71.43	60.50	70.56
SNPEp-2	$\delta_{\text{base}} = 2.0, \beta = 0.22$	87.49	---	---	---

4 总结

本文基于 shape index 对于图像局部二维结构的描述能力,结合尺度空间理论和基于灰度信息空间分解策略提出了一种新的适合于 HEp-2 细胞染色模式分类的方法.由于同时具备对二维图像结构和空间信息的描述能力,本文方法在 ICPR 2012 竞赛数据集和 SNP HEp-2 数据集上获得了非常好的分类效果,其分类性能优于 LBP、CLBP 等纹理特征描述技术,在 ICPR 测试数据集中达到了和竞赛第一名 CoALBP 相当的准确率.

参考文献

- [1] Ponomarev G V, Arlazarov V L, Gelfand M S, et al. ANA HEp-2 cells image classification using number, size, shape and localization of targeted cell regions [J]. Pattern Recognition, 2014, 47(7): 2360 - 2366.
- [2] Stoklasa R, Majtner T, Svoboda D. Efficient k -NN based HEp-2 cells classifier [J]. Pattern Recognition, 2014, 47(7): 2409 - 2418.
- [3] Ul Islam I, Di Cataldo S, Bottino A G, et al. Classification of HEp-2 staining patterns in immune fluorescence images [A]. International Conference on Bioinformatics Models, Methods and Algorithms [C]. Barcelona, Spain: Springer, 2013. 1 - 9.
- [4] Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns [J]. Pattern Recognition, 2009, 42(3): 425 - 436.
- [5] Nosaka R, Suryanto C H, Fukui K. Rotation invariant co-occurrence among adjacent LBPs [A]. Computer Vision-ACCV 2012 Workshops [C]. Berlin: Springer Heidelberg, 2013. 15 - 25.
- [6] Nosaka R, Ohkawa Y, Fukui K. Feature extraction based on co-occurrence of adjacent local binary patterns [A]. Advances in Image and Video Technology [C]. Berlin: Springer Heidelberg, 2012. 82 - 91.

vances in Image and Video Technology [C]. Berlin: Springer Heidelberg, 2012. 82 - 91.

- [7] Nosaka R, Fukui K. Hep-2 cell classification using rotation invariant co-occurrence among local binary patterns [J]. Pattern Recognition, 2014, 47(7): 2428 - 2436.
- [8] Theodorakopoulos I, Kastaniotis D, Economou G, et al. Hep-2 cells classification via fusion of morphological and textural features [A]. Bioinformatics & Bioengineering (BIBE) [C]. USA: IEEE Press, 2012. 689 - 694.
- [9] Di Cataldo S, Bottino A, Islam I U, et al. Subclass discriminate analysis of morphological and textural features for HEp-2 staining pattern classification [J]. Pattern Recognition, 2014, 47(7): 2389 - 2399.
- [10] Koenderink J J, van Doorn A J. Surface shape and curvature scales [J]. Image and Vision Computing, 1992, 10(8): 557 - 564.
- [11] 王永明, 王贵锦. 图像局部不变性特征与描述 [M]. 北京: 国防工业出版社, 2010.
- [12] Foggia P, Percannella G, Soda P, et al. Benchmarking HEp-2 cells classification methods [J]. IEEE Transactions on Medical Imaging, 2013, 32(10): 1878 - 1889.
- [13] Wiliem A, Wong Y, Sanderson C, et al. Classification of human epithelial type 2 cell indirect immune fluorescence images via codebook based descriptors [A]. Applications of Computer Vision (WACV) [C]. USA: IEEE Press, 2013. 95 - 102.
- [14] Chang C C, Lin C J. LIBSVM: A library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 389 - 396.

作者简介



张东波 (通信作者) 男, 1973 年生于湖南邵阳, 现为湘潭大学教授、硕士生导师. 主要研究方向为模式识别、图像处理和机器学习.
E-mail: zhadonbo@163.com



颜霜 女, 1989 年生于湖南湘潭, 现为湘潭大学信息工程学院研究生. 主要研究方向为图像处理与模式识别.
E-mail: yanshuangxu@163.com