

基于视觉词袋模型的亚损伤红细胞识别

郑康,袁瑜含,谷雪莲,鲍睿,郑钰,杨玉菊
上海理工大学医疗器械与食品学院,上海 200093

【摘要】根据亚损伤红细胞的形态学变化,提出一种自动识别亚损伤红细胞的方法。采用体外循环过程中的血细胞图像,包括2763张亚损伤红细胞图像和2507张正常红细胞,利用视觉词袋作为红细胞特征提取方法,分别选用多项核、高斯核、Sigmoid核函数的支持向量机模型。采用5折交叉验证方法验证方法的性能,并选取精确度、召回率、F1评分作为评价指标。结果表明3种不同内核模型的识别准确率分别为 $91.05\%\pm 0.82\%$ 、 $94.16\%\pm 0.50\%$ 、 $85.60\%\pm 0.94\%$ 。本研究提出的方法有效区别了亚损伤红细胞,为亚致死性损伤检测提供自动化方案。

【关键词】亚损伤红细胞;视觉词袋模型;支持向量机;自动识别

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2022)04-0469-06

Identification of sublethally damaged red blood cells based on bag-of-visual words model

ZHENG Kang, YUAN Yuhan, GU Xuelian, BAO Rui, ZHENG Yu, YANG Yuju

School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Abstract: A method for the automatic identification of sublethally damaged red blood cells (RBC) according to their morphological changes is proposed. The images of blood cells during cardiopulmonary bypass are analyzed in the study, including 2763 images of sublethally damaged RBC and 2507 images of normal RBC. The bag-of-visual words is used as the feature extraction method of RBC, and the support vector machine models of polynomial kernel, Gaussian kernel and Sigmoid kernel function are selected separately. A 5-fold cross validation method is adopted to verify the performance of the method, taking precision rate, recall rate and F1 score as evaluation indexes. The results show that the identification accuracies under 3 different kernel models are $91.05\%\pm 0.82\%$, $94.16\%\pm 0.50\%$ and $85.60\%\pm 0.94\%$, respectively. The proposed method can effectively distinguish the sublethally damaged RBC and provide a scheme for the automatic identification of sublethally damaged RBC.

Keywords: sublethally damaged red blood cell; bag-of-visual words model; support vector machine; automatic identification

前言

亚致死性损伤是指在体外循环中各类非生理条件使红细胞发生的非溶血性损伤,即亚损伤^[1-2]。当血液流经机械泵、人工肺、超滤器及插管等循环系统时,血液持续处在湍流流场和剪切力的区域,红细胞受到撞击、剪切和负压的影响,但未导致其直接破裂^[3]。受到亚损伤的红细胞在特定情况下能恢复正常的形态和功能,这类细胞被称为亚损伤红细胞^[4]。

有研究表明亚损伤红细胞的机械脆性和渗透脆性会增加,流变性能和形态学也会发生改变,而且这种变化会引起无法预知的延迟性溶血^[5-7]。但以往血液损伤是直接通过测定血浆中游离血红蛋白溶度来表征溶血率^[8],这种检测方法主要运用于接触血液的生物医学仪器造成的即时性溶血^[9];且未有资料表明亚损伤带来的延迟性溶血可通过游离血红蛋白的指标来量化和检测^[10]。同时有研究表明术后24 h延迟性溶血量可达即时性溶血的5倍^[11-12]。通过对比体外循环手术前和手术后红细胞电镜图可以发现亚损伤红细胞形态发生明显改变,出现棘型、棘球型等各种异常形态的变化^[10, 13-14],由此提出可通过检测红细胞的亚损伤评价循环设备^[10, 15]。研究自动识别亚损伤红细胞的方法对推进血液微观的亚损伤以及临床血液的保护具有重要的意义。

与正常红细胞相比,亚损伤红细胞在形态上主

【收稿日期】2021-10-24

【基金项目】上海市生物医学工程研究生示范实践基地(1017308011)

【作者简介】郑康,硕士在读,研究方向:机器学习、生物医学图像, E-mail: Zachary_k@163.com

【通信作者】谷雪莲,博士,副教授,研究方向:微创外科技术与器械研发, E-mail: guxuelian@usst.edu.cn

要表现为棘形和褶皱;当损伤严重时会出现不同程度的收缩(图1)。为自动化识别出此类细胞,需对细胞膜的特点进行特征提取。现已提出基于图算法的视觉特征提取方法^[16],包括视觉词袋^[17]、哈希算法^[18]等,并且在生物医学研究中取得广泛应用。Kallipolitis等^[19]开发了基于视觉词汇的方法自动对特征进行学习,并对组织病理图片进行区域测量分类;Nayak等^[18]提出利用视觉特征和语义属性以及上下文知识自动分类组织病理图片;Díaz等^[20]提出在

图像中选择和描述局部病理斑块的方法可以提取潜在的语义信息;Raza等^[21]分析了图像中的旋转和尺度不变性问题,发现在肾细胞癌变图像中可以获得更好的分类精度;Bouida等^[22]通过提取大量纹理特征,并在生物医学图像上进行分析 and 测试,发现在生物医学图像中可采用纹理特征对医学图像进行特征提取。本研究采用基于视觉词袋模型方法对亚损伤红细胞进行识别,通过选取分类器的3种内核模型验证特征提取的有效性和方法的可行性。

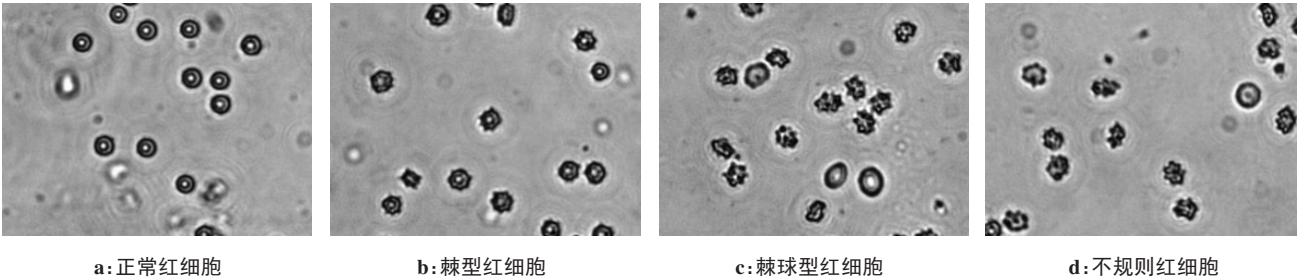


图1 正常红细胞和亚损伤红细胞形态学对比图
Figure 1 Morphological comparison of normal and sublethally damaged red blood cells (RBC)

1 材料和方法

1.1 数据图像采集

本研究的血液图像采集于体外循环中健康的新鲜猪血,实验过程完全按照体外循环标准。实验前,测定猪血中游离血红蛋白和红细胞比容分别为 $(1.32\pm0.03)\%$ 和 $(32.08\pm0.04)\%$,低于溶血指标,满足实验标准。实验过程中,按固定间隔时间由倒置显微镜(徕卡 DMI8)进行拍摄,每幅图像大小为 $1\,920\times2\,560$ 。通过细胞分水岭分割^[23],得到2 507张正常红细胞和2 763张亚损伤红细胞。

1.2 视觉词袋模型

视觉词袋方法是一类利用提取图像中类似特征的分类方法^[24]。如图2所示,算法流程分为4个阶段:(1)使用特征提取方法提取纹理特征或关键点;(2)将关键点聚类生成视觉词汇;(3)将每个视觉词汇编码视觉词袋直方图;(4)使用直方图和图像标签训练分类器。最后通过测试集利用训练集提供的分类标签来预测检测图像的标签。

1.2.1 特征提取方法 图像局部特征描述在目标识别和图像分类领域得到广泛应用,能较好地描述图像纹理。本研究中的模型采用尺度不变特征变换 (Scale-Invariant Feature Transform, SIFT)计算和描述得到图像的关键点和对应的纹理特征^[25],其中每个

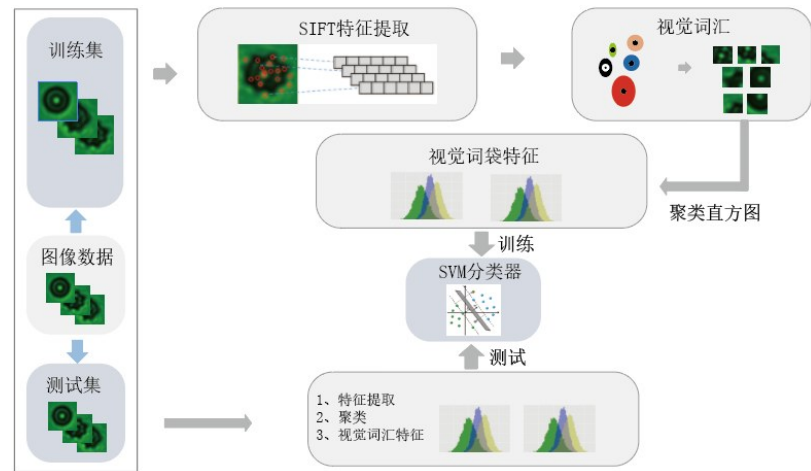
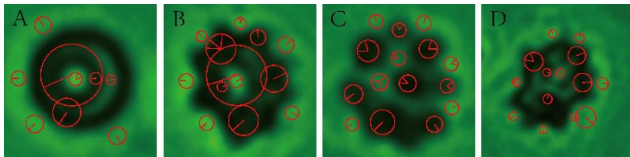


图2 基于视觉词袋模型的亚损伤红细胞识别流程
Figure 2 Identification of sublethally damaged RBC based on bag-of-visual words model

细胞块都具有不同尺度下的纹理特征。

采用此方法的优势是当细胞旋转或者尺度变化时,提取的特征仍保持不变,具有较好的稳定性^[26-27]。特征提取的具体流程是先将图片中每幅图转为灰度图像,然后利用SIFT检测器从细胞形态上去寻找关键点,最后计算每个关键点的纹理特征向量。假设 S 是细胞纹理特征的集合,可定义为 $S = [F_1, F_2, \dots, F_l]^T$,其中 F_l 是第 l 个关键点的纹理特征向量,向量维度为128。

与正常红细胞提取的关键点特征相比,亚损伤红细胞的特征基本上都在细胞膜具有突起和褶皱变化的区域,如图3所示。其中,红色圆圈表示关键点选取的纹理特征范围,圈中直线指向的方法为图像局部的梯度方向,在这个方向上表示关键点对应的纹理特征向量。对比图3中4种红细胞,在图3a、b中该方法均能提取中心圆的纹理特征,而图3b~d中明显在细胞的棘型边缘提取到纹理的梯度变化。当红细胞维持双圆饼型时,SITF可有效提取中心圆特征,同时出现突起和褶皱时,可提取其细胞膜的变化。



a: 正常红细胞 b: 棘型红细胞 c: 棘球型红细胞 d: 不规则红细胞

图3 SIFT提取的关键点

Figure 3 Key points extracted by scale-invariant feature transform

1.2.2 构造视觉词汇 纹理特征提取过程中红细胞图像之间可能具有相同或者类似的纹理特征,当所有纹理特征输入模型会导致模型训练参数过大,模型收敛性和泛化能力较差。通过聚类方法可将所有特征点聚类为 k 个簇,使相似的纹理特征聚集,降低簇间的相似度。然后采用K-means聚类算法对纹理特征向量 F_l 聚类,将聚类返回的聚类中心表示为构造的视觉词汇 $V = \{V_1, V_2, \dots, V_j\}$, $j = 1, 2, \dots, k$ 。其中,每个纹理特征向量 F_l 通过采用欧式距离找到与之最近的视觉词汇,并归属于对应一类簇。

1.2.3 视觉词袋特征提取 通过聚类后得到的 k 个视觉词汇组成的词典称之为词袋。将每张红细胞图片的关键点的纹理特征向量对视觉词汇归类处理的过程称为投袋,原理就是对归为一簇的关键点统计构成直方图,可以表示为 $X = \{x_1, x_2, \dots, x_j\}$, $j = 1, 2, \dots, k$,相对应的视觉词袋特征为:

$$x_j = \sum_{l=1}^N \mu_j(l), j = 1, 2, \dots, k \quad (1)$$

$$\mu_j(l) = \begin{cases} 1, & \|V_j - F_l\| \leq \|V - F_l\|, l = 1, 2, \dots, m \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中, m 为一幅图像的纹理特征向量数量。当纹理特征 F_l 靠近视觉词汇 V_j 时, $\mu_j(l)$ 为1,累计靠近视觉词汇的数量表示词汇特征数量的大小。得到的 X 矩阵代表红细胞提取的视觉词袋特征。

1.2.4 支持向量机 (Support Vector Machine, SVM) 图像分类器 SVM是统计学习理论的发展,试图在空间中找到一个最优的超平面作为决策函数。对于一个两分类识别问题,SVM从无限多的线性决策边界中选择泛化误差最小的决策边界^[28]。因此,所选择的决策边界(用特征空间中的超平面表示)是两类之间剩余边界最大的边界,其中边界定义为两类中最接近的情况到超平面的距离之和。边际距离最大化的问题可以用对偶问题来解决。

为了使细胞的词袋特征作为的特征空间线性可分,利用SVM的核函数技巧将词袋特征映射到高维空间^[29]。给定有标签的数据集 $D = \{(X_i, y_i)\}_{i=1}^m$,其中 x_i 是第 i 张图片的词袋特征向量,对应标签 y_i 为 $\{+1, -1\}$,分别对应亚损伤红细胞和正常红细胞。首先建立线性SVM模型的优化问题^[30]:

$$\min_{w, b, \xi_i} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \right) \quad \text{s.t. } y_i (w^T X_i + b) \geq 1 - \xi_i, \xi_i > 0, i = 1, 2, \dots, m \quad (3)$$

其中, C 为惩罚因子; ξ_i 是松弛因子; $w^T = (w_1; w_2; \dots; w_i)^T$ 为超平面的法向量,决定了超平面的方向; b 表示位移项,决定了超平面与原点之间的距离。

建立划分超平面所对应的模型公式:

$$f(X) = w^T X_i + b \quad (4)$$

令 $\phi(X_i)$ 表示 x_i 映射后的特征向量,其在特征空间的内积 $\phi(X_i)^T \phi(X_j)$ 就等于它们在原始的样本空间的核函数^[31],记为 $K(X_i, X_j)$ 。这里选取了3种核函数,公式见表1。

1.3 评估方法

1.3.1 3类内核模型验证方法

表1 SVM内核表达式

Table 1 Expressions of SVM kernels

名称	表达式
多项式核函数	$K(X_i, X_j) = (X_i^T X_j)^d$
高斯核函数	$K(X_i, X_j) = \exp\left(\frac{\ X_i - X_j\ ^2}{-2\sigma^2}\right)$
Sigmoid核函数	$K(X_i, X_j) = \tanh(\beta X_i^T X_j + \theta)$

分类模型性能的有效方法,常被用来检测模型的泛化能力。首先,红细胞总数被分成K个数目相等的非重叠子集,同时保持每个子集在数据分布上的一致性;然后,在每一次验证中,选择其中一个子集作为验证集,其他子集作为训练集,循环K次;最后,计算K折的平均验证作为最终预测分数。

使用K折交叉验证的主要优点在验证过程中每次输入图像都是从数据集中随机划分选取,有效避免样本训练和测试时出现不准确和不稳定的现象。因此,为了评估模型的总体性能,本研究对数据集进行5折交叉验证。

1.3.2 3类内核模型评价指标 分类器的混淆矩阵表明了正确分类的数量以及基于一组示例在每种情况下真实类别和预测类别的划分情况。因此可通过混淆矩阵找到每个样本的正确分类和错误分类图像的数量信息。对于所提出的评价工具,本研究还使用了精确度、召回率以及F1评分,这些指标是根据混淆矩阵中真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)计算出来的。其中,精确度(Precision)是指对已给的正样本被正确划分的情况,召回率(Recall)是指对所有样本被正确划分的情况,也称为查全率,而F1评分是平衡精确度和召回率一对矛盾的度量,具体计算方法见式(5~7):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

2 结果与分析

2.1 3种内核模型精确度验证

多项核、高斯核、Sigmoid核函数SVM模型在不同验证子集的验证精度上保持稳定,但高斯核函数在模型稳定性和精确度表现最优,其平均精确度为94.16%±0.50%,其次为多项核函数和Sigmoid核函数模型,平均精确度分别为91.05%±0.82%、85.60%±0.94%(图4)。前3折验证子集中,3种内核模型评分均保持较好的稳定性;但在第4和第5验证子集中,高斯核与其他内核模型性能表现上相反。

2.2 3种内核模型召回率验证

3种内核模型5折交叉验证召回率如图5所示。多项核、高斯核、Sigmoid核函数SVM模型召回率分别为90.52%±0.81%、94.08%±0.02%、84.55%±0.96%。其中采用高斯核模型召回率评分保持最优;采用高斯核和多项核模型在不同验证子集评分均保持较好的稳定性;Sigmoid核函数召回率波动性较大。

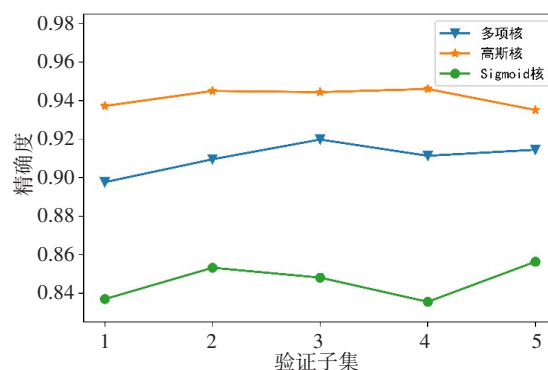


图4 3种内核模型5折交叉验证精确度

Figure 4 Precision rates of 3 kernel models in 5-fold cross validation

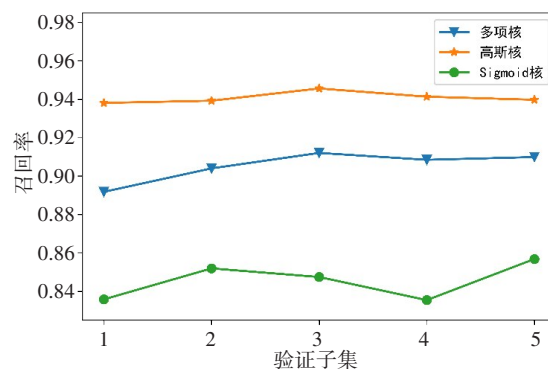


图5 3种内核模型5折交叉验证召回率

Figure 5 Recall rates of 3 kernel models in 5-fold cross validation

2.3 3种内核模型F1评分

多项核、高斯核、Sigmoid核函数SVM模型F1评分分别为90.60%±0.93%、93.73%±0.61%、84.65%±0.95%(图6)。采用高斯核和多项核模型对前3验证子集的数据表现趋于一致,而多项核的趋势更大,但在后2子集中保持相反的趋势,可见在词袋特征映射过程中提取的特征点具有差异性,影响不同内核空间划分的性能。同时,细胞类型在不同子集的数量和形状的差异性,对于正常细胞和亚损伤细胞选择作为正类的敏感性不同。

2.4 分类器混淆矩阵评价

混淆矩阵最能直接显示模型对样本正确和错误识别的数量上的差异性。图7中,纵坐标为样本的真实标签,横坐标为样本的预测标签,其中1代表亚损伤红细胞,作为正类,而-1表示正常红细胞,作为负类,颜色深浅表示数量大小,可见右侧的标尺。

在亚损伤红细胞识别中,多项核、高斯核、Sigmoid核函数SVM模型正确识别数量分别为738、805、738,其中采用高斯核的识别错误数量最少。而在正常红细胞识别中,3种内核模型正确识别数量分别为701、686、625。对比3种内核,高斯核在综合识别数量高于另外两种,并且识别目标作为亚损伤红细胞误差数量最小。

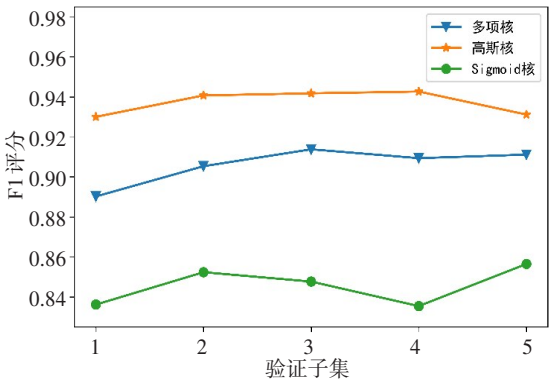


图6 3种内核模型5折交叉验证F1评分

Figure 6 F1-scores of 3 kernel models in 5-fold cross validation

3 讨论

本研究建立了针对亚损伤红细胞特点的视觉词袋模型,在模型分类器部分选用SVM的3种内核做性能进行对比。3种内核模型精确度验证率均在80%以上,表明视觉词袋模型能有效地提取亚损伤红细胞在形态学上的局部特征集合作为分类器的特征表达。

在特征分析中,二次多项式低维度下特征点的线性相关性以及二次多项核SVM模型精确度结果表明由K-means聚类的视觉词袋能映射亚损伤红细胞与正常红细胞形态上的差异性。在3种内核模型的混淆举证中,高斯核函数SVM模型对亚损伤红细胞正确识别

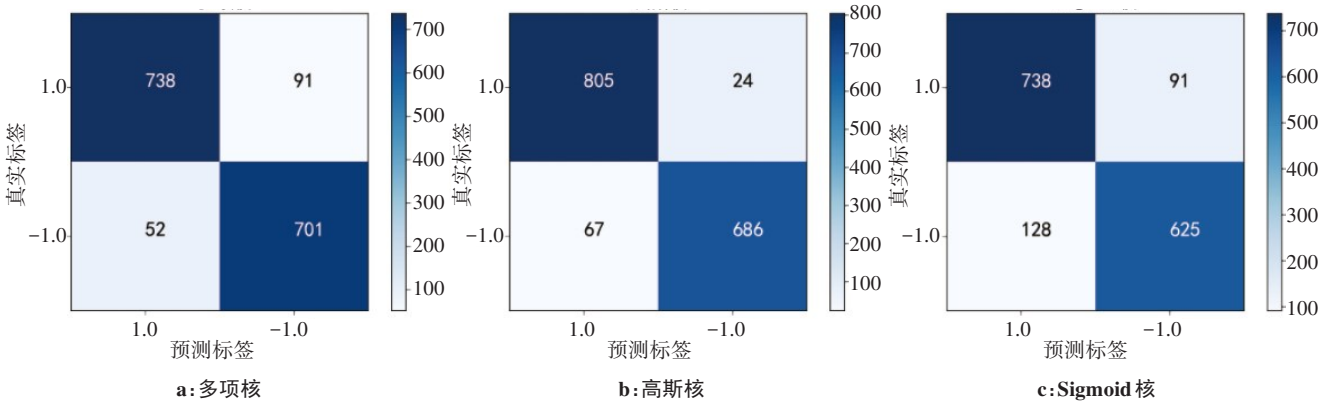


图7 3种内核模型混淆矩阵

Figure 7 Confusion matrixes of 3 kernel models

数量最多,正常红细胞识别数量较少,但亚损伤红细胞识别数量较少的多项核SVM模型对正常红细胞识别数量最多,这种差异表明特征点之间仍存在相关性或者无法作为识别的特征点。同时,结合高斯核内核的特点,部分特征不能在高维空间进行划分,所以这类特征在词汇构建的过程中聚集了两类细胞共同的形态特征信息。在精确度评分中,当验证子集数据改变时,高斯核和其他两种内核模型趋势变化相反,部分词袋特征影响模型的识别过程。因此,高斯核SVM模型在细胞边缘的信息提取仍可以继续改善,如通过特征筛选优化词袋特征的表达过程。另一方面,在红细胞刚发生亚损伤时,细胞表面出现较少的短穗状的针状物,这类特征提取后容易映射到正常红细胞聚类簇的区域。

对比5折交叉验证模型性能在精确度、召回率的趋势,高斯核函数在性能打分以及性能稳定性上均表现最优,体现出高斯核的映射能力更强,能更优地将特征提取的数据映射到高维度;另一方面,高斯核模型的性能标准差均最小,所以采用该内核的泛化性能最优。因此,在视觉词袋模型分类器中应选高斯核SVM分类器;同时Bonaccorso^[31]指出精确率和召回率衡量模型时往往互相矛盾,可通过F1评分以及混淆矩阵综合评定模型性能。当亚损伤细胞作为正类需要识别时,模型可采用

高斯核,当需要识别正常红细胞的情况下,可选用多项核作为模型的内核。

为了体现视觉词袋特征在不同分类器上空间划分的性能,本研究选取常见的机器学习分类模型与高斯核SVM进行对比(表2)。在正常红细胞识别中,精确度最高的为随机森林模型的94%,召回率最高的为高斯核SVM模型的97%;在亚损伤红细胞识别中,精确度最高的为高斯核SVM模型的97%,召回率最高的为随机森林模型的94%。数据证明召回率与精确度之间的矛盾性,但高斯核SVM模型对两类细胞的F1评分均高于其他分类器,因此可通过该方法对亚损伤红细胞进行识别,这与Patle等^[29]的研究结果相一致。

早期Black等^[14]发现血液经由血细胞保存器后,亚损伤红细胞数量显著增加,并提出这些形态变化的红细胞可作为亚损伤的指标。同时Tsui等^[32]通过对红细胞进行显微镜观察和血红蛋白检测,分析得出亚损伤红细胞数量和血红蛋白含量呈正相关。因此,有望利用视觉词袋模型在手术过程中对红细胞形态变化进行识别,从而检测血液接触医疗器械对血液的亚损伤变化情况的定性分析。另一方面,亚损伤红细胞相关形态学研究仍具有局限性,未有研究细分亚损伤红细胞的不同亚型,导致在血液亚损伤上难以进行定量研究。

表 2 不同分类器的模型分类精度(%)
Table 2 Model classification accuracies of different classifiers (%)

模型	正常红细胞			亚损伤红细胞		
	精确度	召回率	F1 评分	精确度	召回率	F1 评分
高斯核 SVM	91	97	94	97	91	94
装袋算法 ^[33]	89	90	89	91	90	90
随机森林 ^[34]	94	89	91	91	94	93
决策树 ^[35]	90	79	84	83	93	87
集成学习 ^[36]	91	94	93	93	90	92

4 结 论

本研究根据血液在发生亚损伤过程中,红细胞发生形态学变化,提出一种基于视觉词袋模型的识别方法。其中,SITF 能有效提取正常红细胞和亚损伤红细胞在细胞膜上的形态学关键点,构建的视觉词汇特征能表达亚损伤细胞的视觉信息,并通过对不同内核的性能,确定高斯核的性能最优,最后对不同分类器的差异。使用本研究提出的方法能有效地分类出亚损伤红细胞,为推动亚致死性损伤检测提供自动化的方案。

【参考文献】

[1] POLASCHEGG H D, LEVIN N W. Hemodialysis machines and monitors[M]//Replacement of renal function by dialysis. Dordrecht: Springer Netherlands, 2004: 325-449.

[2] TANG H L, HANKA R, IP H H. Histological image retrieval based on semantic content analysis[J]. IEEE Trans Inform Technol Biomed, 2003, 7(1): 26-36.

[3] SAYLOR D M, BUEHLER P W, BROWN R P, et al. Predicting plasma free hemoglobin levels in patients due to medical device-related hemolysis[J]. ASAIO J, 2019, 65(3): 207-218.

[4] MCNAMEE A P, RICHARDSON K, HOROBIN J, et al. Susceptibility of density-fractionated erythrocytes to subhaemolytic mechanical shear stress[J]. Int J Artif Organs, 2019, 42(3): 151-157.

[5] HASHIMOTO S. Detect of sublethal damage with cyclic deformation of erythrocyte in shear flow[J]. J Syst Cybern Inform, 2014, 12(3): 41-46.

[6] GRIGIONI M, DANIELE C, MORBIDUCCI U, et al. Potential mechanical blood trauma in vascular access devices: a comparison of case studies[J]. Int J Artif Organs, 2002, 25(9): 882-891.

[7] JAMES M E, PAPAVALASSILIOU D V, O'REAR E A. Use of computational fluid dynamics to analyze blood flow, hemolysis and sublethal damage to red blood cells in a bileaflet artificial heart valve[J]. Fluids, 2019, 4(1): 19.

[8] OLIA S E, MAUL T M, ANTAKI J F, et al. Mechanical blood trauma in assisted circulation: sublethal RBC damage preceding hemolysis[J]. Int J Artif Organs, 2016, 39(4): 150-159.

[9] DENG L, YANG H, MENG W, et al. Ultrastructural changes in red blood cells in aortic dissection patients under extracorporeal circulation: atomic force microscopy study[J]. Int J Artif Organs, 2018, 41(8): 452-459.

[10] REINHART W H. Peculiar red cell shapes: fahraeus Lecture 2011[J]. Clin Hemorheol Microcirc, 2011, 49(1-4): 11-27.

[11] BASKURT O K, UYUKLU M, MEISELMAN H J. Protection of erythrocytes from sub-hemolytic mechanical damage by nitric oxide mediated inhibition of potassium leakage[J]. Biorheology, 2004, 41: 79-89.

[12] CHEN G L, QIAO Y R, MA J H, et al. Extracorporeal cardiopulmonary resuscitation in children of Asia pacific: a retrospective analysis of extracorporeal life support organization registry[J]. Chinese Medical Journal, 2018, 131(12): 1436-1443.

[13] ZENI C, BOVOLENTA M R, STAGNI A. Occurrence of echinocytosis in circulating RBC of black bullhead, Ictalurus melas (Rafinesque), following exposure to an anionic detergent at sublethal concentrations[J]. Aquat Toxicol, 2002, 57(4): 217-224.

[14] BLACK K L, JONES R D. The discocyte-echinocyte transformation

as an index of human red cell trauma[J]. Ohio J Sci, 1976, 76(5): 225-230.

[15] BASAVANHALLY A N, GANESAN S, AGNER S, et al. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology[J]. IEEE Trans Biomed Eng, 2010, 57(3): 642-653.

[16] CAICEDO J C, CRUZ A, GONZALEZ F A. Histopathology image classification using bag of features and kernel functions [C]. Artificial Intelligence in Medicine, 2009: 126-135.

[17] ZHANG X, LIU W, DUNDAR M, et al. Towards large-scale histopathological image analysis: hashing-based image retrieval[J]. IEEE Trans Med Imaging, 2015, 34(2): 496-506.

[18] NAYAK N, CHANG H, BOROWSKY A, et al. Classification of tumor histopathology via sparse feature learning[C]. 2013 IEEE 10th International Symposium on Biomedical Imaging, San Francisco: IEEE, 2013: 410-413.

[19] KALLIPOLITIS A, MAGLOGIANNIS I. Creating visual vocabularies for the retrieval and classification of histopathology images [C]. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2019: 7036-7039.

[20] DÍAZ G, ROMERO E. Histopathological image classification using stain component features on a pLSA model[C]. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Berlin: Springer, 2010: 55-62.

[21] RAZA S H, PARRY R M, MOFFITT R A, et al. An analysis of scale and rotation invariance in the bag-of-features method for histopathological image classification[C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011, Berlin: Springer, 2011: 66-74.

[22] BOUIDA A, BELADGHAM M, BASSOU A, et al. Evaluation of textural degradation in compressed medical and biometric images by analyzing image texture features and edges[J]. Traitement du Signal, 2020, 37(5): 753-762.

[23] JIANG P, ZHANG X, WANG F. Red blood cell detection by the improved two-layer watershed segmentation method with a full convolutional neural network[J]. J Med Imag Health Inform, 2018, 8(1): 50-54.

[24] TIAN L, WANG S. Improved bag-of-words model for person re-identification[J]. Tsinghua Science and Technology, 2018, 23(2): 145-156.

[25] PAL R, SARASWAT M, MITTAL H. Improved bag-of-features using grey relational analysis for classification of histology images[J]. Complex Intell Syst, 2021, 7(3): 1429-1443.

[26] GUPTA S, THAKUR K, KUMAR M. 2D-human face recognition using SIFT and SURF descriptors of face's feature regions[J]. Visual Comput, 2020: 1-10.

[27] TAREEN S A, SALEEM Z. A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK [C]//2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). Sukkur: IEEE, 2018: 1-10.

[28] NOBLE W S. What is a support vector machine?[J]. Nat Biotechnol, 2006, 24(12): 1565-1567.

[29] PATLE A, CHOUHAN D S. SVM kernel functions for classification [C]. 2013 International Conference on Advances in Technology and Engineering (ICATE). Phuket: IEEE, 2013: 1-9.

[30] MIRANDA P B, PRUDENCIO R B, DE CARVALHO A P, et al. A hybrid meta-learning architecture for multi-objective optimization of SVM parameters[J]. Neurocomputing, 2014, 143: 27-43.

[31] BONACCORSO G. Machine learning algorithms[M]. Birmingham: Packt Publishing Ltd, 2017.

[32] TSUI S M, AHMED R, AMJAD N, et al. Single red blood cell analysis reveals elevated hemoglobin in poikilocytes[J]. J Biomed Opt, 2020, 25(1): 1-13.

[33] LIU Y, LONG F. Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning [C]//ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Singapore: Springer, 2019: 113-121.

[34] SARASWAT M, ARYA K V. Feature selection and classification of leukocytes using random forest[J]. Med Biol Eng Comput, 2014, 52 (12): 1041-1052.

[35] TANTIKITTI S, TUMSWADI S, PREMCHAIWADI W. Image processing for detection of dengue virus based on WBC classification and decision tree[C]//2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015). Bangkok: IEEE, 2015: 84-89.

[36] TSUJI K, LU H, TAN J K, et al. Automatic identification of circulating tumor cells in fluorescence microscopy images based on adaBoost[C]// 2017 17th International Conference on Control, Automation and Systems (ICCAS). Jeju: IEEE, 2017: 1449-1454.

(编辑:谭斯允)