

DOI: 10.13382/j.jemi.B2205178

基于谱分解动态散射成像的细胞无标记 检测与分类方法*

周雨宁¹ 夏华² 王晓莉³ 柳竞涵² 翟嘉¹ 李晨曦² 蒋景英^{1,4}

(1. 天津大学医学工程与转化医学研究院 天津 300072; 2. 天津市生物医学检测技术与仪器重点实验室 天津 300072;
3. 潍坊医学院 山东 261053; 4. 北京航空航天大学医学科学与工程学院 北京 100091)

摘要:细胞成像及检测技术在医学研究及临床诊断领域具有重要的研究意义和应用价值,而无标记与高通量检测尤其具有挑战。本研究基于动态散射理论的细胞成像方法,搭建了动态散射成像系统,提出了基于谱分解的动态信号提取算法,结合机器学习算法实现了无标记、高通量的细胞分类。采用血细胞、EG7-OVA 肿瘤细胞、A549 肺癌肿瘤细胞进行实验验证,结果表明本文提出的方法对血细胞与肿瘤细胞识别准确率可达98%以上,对于血细胞、EG7-OVA 细胞和 A549 细胞之间的三分类识别率约为91%。本文实现的细胞检测和分类方法具有临床应用前景。

关键词:动态散射成像;谱分解;无标记细胞检测与分类;机器学习

中图分类号: O436.2 **文献标识码:** A **国家标准学科分类代码:** 140.3060

Study on label-free cell detection and classification method by using spectral decomposition-based dynamic scattering imaging

Zhou Yuning¹ Xia Hua² Wang Xiaoli³ Liu Jinghan² Zhai Jia¹ Li Chenxi² Jiang Jingying^{1,4}

(1. Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, China; 2. Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments, Tianjin University, Tianjin 300072, China; 3. Weifang Medical University, Shandong 261053, China; 4. School of Engineering Medicine, Beihang University, Beijing 100091, China)

Abstract: Cell imaging and detection are of great significance in the field of biomedical research and clinical diagnosis, while label-free and high-throughput detections are particularly challenging. On the basis of dynamic scattering theory, this study built a dynamic scattering imaging system, proposed a spectral decomposition-based dynamic signal extraction algorithm, and achieved label-free and high-throughput cell classification by combining machine learning algorithms. Blood cells, EG7-OVA tumor cells and A549 lung cancer tumor cells are used to verify the current method. Experimental results show 98% accuracy for binary classification of blood cells and tumor cells, and 91% accuracy for the three-type classification of blood cells, EG7-OVA and A549. In summary, the proposed method provides high-throughput, label-free cell detection and classification, and is potential for clinical application.

Keywords: dynamic light scattering imaging; spectral decomposition; label-free cell detection and classification; machine learning

0 引言

细胞是构成人体最基本的结构和功能单位,其数量、形态等特性反映了生理病理状态^[1],是疾病诊断的重要依据^[2]。细胞高通量分析与分选技术也是医学研究及临床诊断热点^[3],在癌症的早期诊断与预后评估等方面具有重要的研究意义和应用价值^[4]。流式细胞分析是临床

常用的细胞检测技术,利用单克隆抗体及荧光探针检测方法实现不同种类细胞计数与特征检测^[5]。该方法对于仪器及检测条件要求较高,并需要对细胞进行标记,检测成本较高,细胞在检测区域内的聚集与重叠也会影响分析结果。共聚焦成像方法可以用于细胞形态分辨,但是其成像视场受限,很难实现高通量检测^[6]。微流控芯片根据细胞尺寸不同可实现无标记捕捉细胞^[7],但是细胞大小及形变能力具有异质性,此方法具有较低的灵敏性

收稿日期: 2022-02-12 Received Date: 2022-02-12

* 基金项目: 国家自然科学基金(81971662, 81871396)、北京市自然科学基金(7202105)、天津市自然科学基金(20JCZDJC00630)项目资助

和特异性,同时需要对不同的细胞设计不同的芯片结构。为了满足临床应用需求,迫切需要开发高通量、无标记、可实现多种细胞同时检测的仪器与方法^[8]。

光散射方法在散射及相干检测理论上,通过分析散射光的空间及时间干涉信号,实现亚微米量级散射特性定量分析,具有无标记、快速无损的优点,在粒度分析、细胞成像等领域应用广泛^[9-11]。相干光照射到细胞后,散射光的空间强度分布会随细胞的尺寸、形状、折射率等变化,形成复杂的干涉图像。干涉信号的时间-空间强度分布及频率变化既反映了细胞形态,又与细胞内部亚微米量级散射特征有关。与荧光标记技术相比,光散射技术无需标记。

根据散射体是否静止,散射光可分为弹性散射光与动态散射光。相关研究证明细胞的弹性散射信号反映了细胞形态^[12]及细胞器分布^[13]。Su等^[14]分析了正常造血细胞和白血病细胞的散射图像,发现线粒体的分布会影响散射光图案中散斑空间分布,为散射光无标记识别不同类型细胞提供了参考;Liu等^[15]使用光学细胞模型模拟单细胞的散射成像,对比3种不同核质比的前列腺肿瘤细胞与正常前列腺细胞散射图像,证明细胞核质比与形成的散射信号具有统计学关系,但未使用真实细胞验证;Su等^[16]使用AdaBoost算法对正常宫颈细胞和HeLa细胞的弹性光散射图案构建了准确率达90%的二分类模型。弹性光散射技术能够实现细胞分类,但是无法实现高通量检测。

当散射体发生运动,产生的散射光不仅仅包含散射体的结构信息,还包含了运动状态的信息。动态光散射由于受背景信号干扰大,目前常用于粒度分析及运动状态测量。Chicea等^[17]在动态光散射时变信号基础上,利用人工神经网络算法评估发酵过程中酵母细胞大小随时间的变化,实现对细胞尺寸的测量。杨斌等^[18]对动态散射信号进行互相关分析,实现对循环流化床内颗粒速度的定量测量。

细胞的动态光散射信号包含了细胞的结构信息与细胞运动信息。在频域上,细胞产生的动态光散射信号区别于背景产生的光散射信号,因此可以通过对信号的频域分析,滤除背景干扰,得到细胞的动态光散射信号。本文提出了基于谱分解的动态信号提取算法,能够从背景光散射信号中提取出运动的细胞光散射信号,提高了成像对比度。进一步,为了提高细胞检测分类效率,需要在细胞动态散射图像特征定量分析的基础上,与机器学习等自动分类算法相结合进行讨论。

本文针对活体细胞无标记成像及分类需要,搭建了动态散射成像系统,研究了基于谱分解的细胞动态散射信号提取方法,提高细胞散射成像对比度及信噪比。通过量化分析细胞内部散射特征纹理参数,并与机器学习

方法相结合,建立细胞分类模型,为高通量细胞成像与检测研究提供了基础。

1 方法

1.1 动态散射成像系统

本文搭建的动态散射成像系统如图1所示,主要由激光器、光束整形装置、显微成像系统组成。激光器采用高稳定氦氖激光器(632.8 nm, 10 mW, MellesGriot),光束整形装置由空间滤波器及准直镜组成。空间滤波器通过物镜与小孔光阑组合去除光束中的高阶模、波面的畸变与噪声,得到更为理想的球面光波,保证光束质量以及空间相干特性。滤波后光束经过焦距为300 mm的准直透镜,均匀照射到细胞成像面,消除了边缘畸变的影响。显微成像系统采用长焦距复消色差显微物镜(10X, WD=33.5 mm, NA=0.28)与高速工业相机(acA800-510uc, 像素尺寸4.8 μm×4.8 μm, 500 fps, Basler),满足动态相干信号检测要求,可实时记录经过细胞散射后形成的动态散射相干信号。

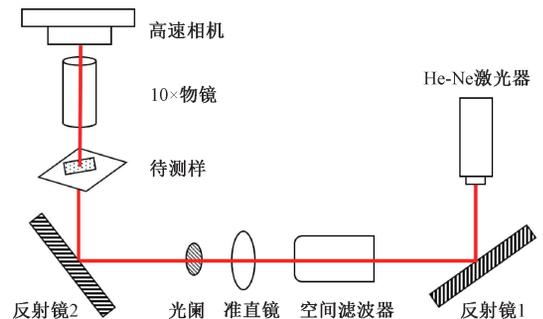


图1 动态散射成像系统示意图

Fig. 1 Diagram of dynamic scattering imaging system

实验样品为血细胞, A549 肺癌肿瘤细胞及 EG7-OVA 肿瘤细胞。实验中采用微流控芯片模拟细胞在血液中流动状态,流道深度为30 μm,宽度为100 μm。微流控系统由微流控芯片、聚四氟乙烯管、微量进样器和注射器泵组成。控制细胞在微通道内的流速约为1 mm/s。为了与细胞流速匹配,图像采集速率设置为400 fps,曝光时间为2 ms。将使用RPMI-1640细胞培养液稀释的血液、A549肺癌肿瘤细胞、EG7-OVA肿瘤细胞通入微流控系统,采集得到3种细胞的动态散射图像。

1.2 基于谱分解的动态散射成像处理算法

为了提高细胞动态散射图像分辨率及信噪比,本文在动态散射理论上,提出了基于动态散射相干信号时频变化特征的滤波处理算法。在信号相关分析及谱分解方法基础上,滤除背景散射影响。其基本原理是假设在成像面上接收到的光散射信号主要包含细胞动态散射

与背景散射所引起的静态散射信号,信号表达式为:

$$I = I_d + I_b \quad (1)$$

其中, I_d 为细胞动态散射信号, I_b 为背景静态散射信号。两种散射信号,两者相互独立。为了从原始信号中区分不同来源的信号,进而滤除背景散射,首先计算原始视频的协方差矩阵,公式如下:

$$\mathbf{R}_x = E(I \times I^T) \quad (2)$$

根据矩阵分解理论,协方差矩阵还可以表示为:

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \quad (3)$$

其中, $\mathbf{U} = [\vec{u}_1, \dots, \vec{u}_j, \dots, \vec{u}_M]$, $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_j, \dots, \lambda_M)$, \vec{u}_j 和 λ_j 分别为协方差矩阵 \mathbf{R} 的第 j 个特征向量和与之对应的特征值。

根据动态光散射理论,不同尺寸及形态的散射体的动态散射信号的空间分布及频率不同,分布在不同特征向量 \vec{u}_j 上所占比重也不同。其中背景散射主要分布在代表低频的特征向量中。在此基础上,设计基于谱分解的滤波算法,在滤除 n 个低频特征向量后,得到细胞动态散射信号,计算公式如下:

$$I_d = \left(\sum_{j=n}^M \vec{u}_j \cdot \vec{u}_j^T \right) I \quad (4)$$

细胞的动态散射信号主要与细胞形态和内部散射结构有关,为细胞分类识别提供了基础。

1.3 细胞动态散射特征参数量化分析

血液由血细胞和血浆构成,血细胞又包括红细胞、血小板和白细胞。通常正常红细胞直径约为 $8 \mu\text{m}$,血小板直径为 $2 \sim 4 \mu\text{m}$,白细胞直径为 $8 \sim 12 \mu\text{m}$,而肿瘤细胞由于异常增殖具有更大的核质比与体积,直径为 $12 \sim 25 \mu\text{m}$ ^[19]。首先对采集到的细胞图像滤除直径小于 $6 \mu\text{m}$ 的细胞图像。进一步,为了提高细胞分类识别准确度及效率,本研究不但计算了细胞基本形态参数,如面积、周长等,并且进一步提取了细胞动态散射图像的纹理特征。采用灰度共生矩阵(GCLM)及其统计特征作为定量指标,分别计算灰度矩阵的同质性(homogeneity)、对比度(contrast)、相异性(dissimilarity)、ASM 能量(ASM energy)、能量(energy)和相关度(correlation),用于表征纹理特征。其中,同质性体现的是灰度值与对角线灰度的接近程度;对比度和相异性反映图像的清晰度以及纹理的深浅;ASM 能量和能量反映图像灰度分布均匀程度以及纹理粗细度;相关度反映图像在行或列上的相似性,各个参数计算公式如下:

$$\text{Homogeneity} = \sum_{i=1}^N \sum_{j=1}^N \frac{\mathbf{g}(i,j)}{1 + (i-j)^2} \quad (5)$$

$$\text{Contrast} = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 \mathbf{g}(i,j) \quad (6)$$

$$\text{Dissimilarity} = \sum_{i=1}^N \sum_{j=1}^N \mathbf{g}(i,j) |i-j| \quad (7)$$

$$\text{ASM} = \sum_{i=1}^N \sum_{j=1}^N \mathbf{g}(i,j)^2 \quad (8)$$

$$\text{Energy} = \sqrt{\text{ASM}} \quad (9)$$

$$\text{Correlation} = \sum_{i=1}^N \sum_{j=1}^N \mathbf{g}(i,j) \left[\frac{(i-\mu_x)(j-\mu_y)}{\sqrt{(\sigma_x^2)(\sigma_y^2)}} \right] \quad (10)$$

其中, $\mathbf{g}(i,j)$ 是样本灰度共生矩阵, N 是灰度共生矩阵的尺寸大小,而 $\mu_x = \sum_i \sum_j i \cdot \mathbf{g}(i,j)$, $\mu_y = \sum_i \sum_j j \cdot \mathbf{g}(i,j)$, $\sigma_x = \sqrt{\sum_i \sum_j (i-\mu_x)^2 \mathbf{g}(i,j)}$, $\sigma_y = \sqrt{\sum_i \sum_j (j-\mu_y)^2 \mathbf{g}(i,j)}$ 。

1.4 分类模型

在细胞动态成像及量化参数基础上,建立基于机器学习算法的分类模型,可实现不同种类细胞识别分类。本文所研究的分类模型主要基于如下 4 种机器学习方法:支持向量机、K-近邻、决策树及随机梯度下降。

支持向量机(support vector machines, SVM)是定义在特征空间上的间隔最大的线性分类器,通过模型训练构建能够使得分类结果之间几何间隔最大的超平面,实现正确分类。该超平面也称为决策面,为了提高决策面的泛化能力与非线性能力,选择合适的核函数是关键。基于线性核函数的 SVM 分类器效率高,对于特征维度较高、样本数目较少分类效果较好。本研究中,针对分类任务中参数多,样品数少的特点,构建基于线性核函数的分类器,可以实现多分类。

K-近邻算法(K-nearest neighbor, KNN)的算法思想是通过测量不同特征值之间的距离进行分类。将待测数据的特征与训练集中的数据进行比较,找到训练集中与其最邻近的一个或者几个样本的类别来决定待测数据的所属类别。

决策树算法(decision tree, ID3)通过建立树形结构,将每个数据从根节点输入,每个叶子节点代表类或者类的分布。通过模型训练,调整每个节点的判断阈值,最终得到决策树模型。测试集的数据通过根节点输入,由当前节点的阈值判断下一层进入某个子树,最后由叶子节点输出类别判断。

随机梯度下降(stochastic gradient descent, SGD)通过设计损失函数,每次随机选取数据来获取梯度下降最快的方向,以此更新损失函数,最后由损失函数判断测试集的类别。

针对小样本分类要求,以上方法在分类模型训练环节采用十折交叉验证方式^[20]。将数据集分成十等份,每次训练取出其中 9 份作为训练集,剩余的一份作为测试

集。在重复 10 次训练后,用 10 次结果的平均值验证模型分类性能。

1.5 分类模型评价

为了评估模型分类效果,分别计算分类结果的准确率 (accuracy)、查准率 (precision)、召回率 (recall)、F1 score 参数。其中,准确率代表的是所有样本中分类正确的比例,查准率代表的是被分类为所需目标的样本中分类正确的比例,召回率代表的是目标样本被正确分类的比例,而 F1 score 是综合查准率与召回率表现的评判指标。各参数计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (11)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (14)$$

其中,TP 是正确分类为阳性的样本数, TN 是正确分类为阴性的样本数, FP 是错误分类为阳性的样本数, FN 是错误分类为阴性的样本数。

2 结果分析

2.1 细胞动态散射图像分析

系统成像得到细胞原始动态散射图像与经过滤波算法处理后的图像如图 2 所示。原始图像中混杂了大量背景散射信号,较难分辨出细胞形态特征。经过基于谱分解的动态光散射成像算法处理后,背景散射被滤除,得到的细胞动态散射图像信噪比较高,可以清楚分辨细胞大小及内部散射特征,为细胞形态及结构参数量化提供了基础。

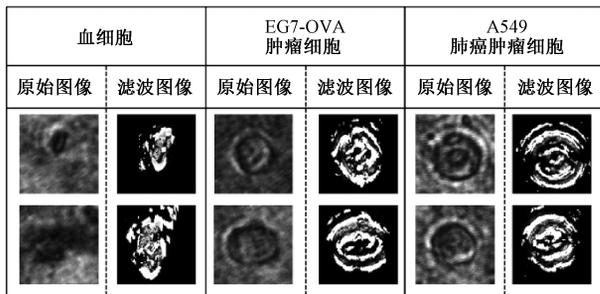


图 2 细胞动态散射图像滤波前后对比

Fig. 2 Comparison of cell dynamic scattering images before and after filtering

从动态散射图像中不仅可以有效分辨细胞大小形

态,其散射纹理特征也反映了细胞内部散射结构信息。大部分血细胞图像与图 2 中展示的第 1 张血细胞图像类似,在显微镜下观察确定为红细胞。红细胞作为一种无核细胞,其散射结构相对简单,与之对应的图像特征表现为红细胞两端的动态散射信号强于其他部分信号。第 2 张血细胞图像为白细胞,散射特征也比较简单,但直径大于红细胞。而癌细胞由于细胞的异常分化,出现细胞核增大、细胞核形状发生改变、核质比例失常现象,因此在动态散射图像中所表现的散射纹理特征更为复杂。

2.2 基于动态散射成像的细胞分类结果

在细胞动态散射成像基础上,提取细胞的形态及动态散射纹理参数作为分类模型的输入参数,分别采用基于 SVM、KNN、ID3 与 SGD 四种方法的模型对 3 种细胞进行分类。比较不同模型分类结果,如表 1 所示。

表 1 模型分类效果评价表

Table 1 Evaluation of model classification

Model	Accuracy	Precision-CTCs	Recall-CTCs	F1 score
SVM	0.911 0	0.996 7	0.889 6	0.994 7
KNN	0.911 2	0.994 0	0.892 6	0.994 2
ID3	0.904 7	0.990 5	0.887 7	0.992 2
SGD	0.832 2	0.962 0	0.822 0	0.975 7

从表 1 中可以看出基于 SVM 和 KNN 方法的分类模型对于 3 种细胞的分类精度均达到 91% 以上,并且查准率,召回率, F1 score 等参数也达到 99%, 88%, 99%, 对于 3 种细胞的分类效果较好。SGD 方法分类精度略差,主要由于该方法不能保证被优化的函数达到全局最优解,受限于损失函数是否为凸函数,对于多参数分类效果较差。SVM 与 KNN 方法对于多参数,小样本的情况具有较好的分类精度。此外,这两种分类模型可以实现非线性分类,为基于细胞多特征参数的识别分类提供基础。为了测试 4 种算法在大样本下的表现情况,使用旋转预处理操作进行数据扩增,并对比了在不同样本量下 4 种算法的表现情况。

从表 2 中可以看出,SGD 算法一直表现不佳,而 ID3 算法的识别准确率异常提升是因为产生的新数据与原数据存在一定相似性,所以 ID3 算法产生了过拟合现象。随着样本量的提升,SVM 和 KNN 方法的识别准确率均有所提高。KNN 方法随着数据扩增识别准确率提升更快,从原理上分析是由于扩增的数据在特征上更接近原数据,而 KNN 是通过选择训练集中特征最接近的类别作为预测类别,因此 KNN 的识别准确率可能在不同种类细胞数量不均匀分布的训练集中产生不稳定的结果。相对来说 SVM 算法对于不同大小的数据集表现更为稳定,鲁棒性更强。

表 2 不同大小数据集下模型分类效果表

Table 2 Accuracy of model classification in data sets of different sizes

样本量	4 000	8 000	12 000	16 000	20 000	24 000
SVM	0.911 0	0.924 9	0.925 8	0.925 9	0.929 5	0.931 4
KNN	0.911 2	0.932 2	0.935 7	0.937 3	0.944 3	0.949 6
ID3	0.904 7	0.978 5	0.984 5	0.986 7	0.988 9	0.990 8
SGD	0.832 2	0.836 2	0.834 7	0.853 2	0.828 1	0.831 0

3 种细胞中 EG7-OVA 与 A549 均为肿瘤细胞,为了进一步比较不同分类模型对于特定细胞种类的分类识别效果,分别比较了 SVM 与 KNN 分类模型的混淆矩阵。利用混淆矩阵统计分析两种模型对于血细胞/肿瘤细胞,以及血细胞/A549 肺癌肿瘤细胞/EG7-OVA 肿瘤的二分类及三分类结果。

考虑到血液中的实际分布,血细胞数量远大于肿瘤细胞数量,制作了血细胞、EG7-OVA 肿瘤细胞、A549 肿瘤细胞数量比例约为 5 : 1 : 1 的数据集。如图 3 所示,SVM 模型针对血细胞/肿瘤细胞的二分类准确率可达 97% 以上,而 KNN 的二分类模型对肿瘤细胞的识别准确率仅为 80%。这是因为当数据集中血细胞的数量远大于肿瘤细胞数量的时候,目标细胞特征的周围出现血细胞的可能性更大,所以 KNN 将很多的 EG7-OVA 肿瘤细胞错误分类成血细胞。对于循环肿瘤细胞检测的实际意义而言,从数量众多的血细胞中识别出肿瘤细胞的临床意义更为重要,因此 SVM 算法更适用于此分类任务。本分类算法对于肿瘤细胞之间的识别仍存在局限性,有 15.93% 的 EG7-OVA 细胞被误判为 A549 细胞,同样也有 11.81% 的 A549 细胞被误判为 EG7-OVA 细胞。原因在于对细胞动态散射特征参数的提取更多地考虑了血细胞与肿瘤细胞之间具有区别的特征。在之后的研究中为了提供不同种类肿瘤细胞分类识别精度,将从细胞特征量化参数选择,成像方法优化等方面进行改进,拓展动态散射成像技术在细胞检测中的应用。

3 结 论

本文针对临床细胞分析需要,研究了基于动态光散射相干成像的活细胞无标记成像及分类识别方法。在动态散射成像系统基础上,开发了基于谱分解的细胞动态散射信号提取方法,提高细胞散射成像对比度及信噪比。为了实现基于细胞散射特征的识别分类,计算细胞内部散射特征纹理参数,并将其作为基于机器学习的分类模型输入参数。设计了多种分类模型,采用血细胞,EG7-OVA 肿瘤细胞,A549 肺癌肿瘤细胞进行细胞成像与分类验证实验。结果表明利用动态散射成像与机器学习算法,可以实现血细胞与肿瘤细胞高精度分类,识别准确率可达 98% 以上,对于不同细胞的三分类精度也可达到

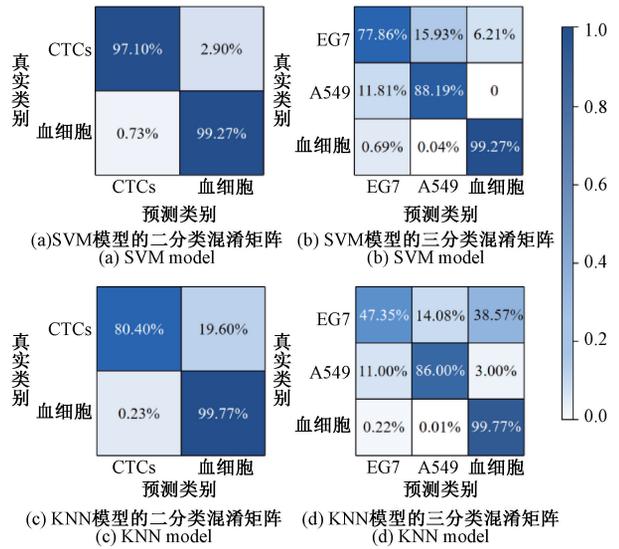


图 3 SVM 和 KNN 模型的混淆矩阵

Fig. 3 Confusion matrixes for SVM and KNN Confusion matrix detailing the classification results of two different cells using

91% 左右。本文研究的系统与方法为高通量、无标记、多种细胞同时检测的临床细胞检测诊断研究提供了参考,具有较好的应用前景。

参考文献

- [1] TRUONG M, FRYE T, MESSING E, et al. Historical and contemporary perspectives on cribriform morphology in prostate cancer [J]. Nature Reviews Urology, 2018, 15(8) : 475-482.
- [2] ALIZADEH E, CASTLE J, QUIRK A, et al. Cellular morphological features are predictive markers of cancer cell state [J]. Computers in Biology and Medicine, 2020, 126: 104044.
- [3] MAIR B, ALDRIDGE P M, ATWAL R S, et al. High-throughput genome-wide phenotypic screening via immunomagnetic cell sorting [J]. Nature Biomedical Engineering, 2019, 3(10) : 796-805.
- [4] LI Y, TIAN X, GAO L, et al. Clinical significance of circulating tumor cells and tumor markers in the diagnosis of lung cancer [J]. Cancer Medicine, 2019, 8(8) : 3782-3792.
- [5] HU L, CHEN X, CHEN M, et al. Enrichment and detection of circulating tumor cells by immunomagnetic beads and flow cytometry [J]. Biotechnology Letters, 2021, 43(1) : 25-34.
- [6] HU Y, TANG W, CHENG P, et al. Monitoring circulating tumor cells in vivo by a confocal microscopy system [J]. Cytometry Part A, 2019, 95(6) : 657-663.
- [7] REN X, FOSTER B M, GHASSEMI P, et al.

- Entrapment of prostate cancer circulating tumor cells with a sequential size-based microfluidic chip[J]. *Analytical Chemistry*, 2018, 90(12): 7526-7534.
- [8] 黄彩虹, 易定容, 金福江, 等. 单细胞分离方法及仪器研究进展[J]. *仪器仪表学报*, 2020, 41(5): 140-153.
HUANG C H, YI D R, JIN F J, et al. Progress on single cell isolation methods and instruments [J]. *Chinese Journal of Scientific Instrument*, 2020, 41(5): 140-153.
- [9] LI Y, NOWAK C M, PHAM U, et al. Cell morphology-based machine learning models for human cell state classification [J]. *NPJ Systems Biology and Applications*, 2021, 7(1): 1-9.
- [10] WAN W Y, LIU L, LIU X, et al. Integration of light scattering with machine learning for label free cell detection [J]. *Biomedical Optics Express*, 2021, 12(6): 3512-3529.
- [11] QUIÑONES R, MORENO S, SHOUP D, et al. Examining particle size of inorganic active ingredients within sunscreens using dynamic light scattering [J]. *Journal of Chemical Education*, 2021, 98 (4): 1371-1380.
- [12] CASAS-RAMOS M A, SANDOVAL-ROMERO G E. Yeast cells characterization through near-forward light scattering[J]. *Optical and Quantum Electronics*, 2020, 52(3): 1-8.
- [13] MARINA O C, SANDERS C K, MOURANT J R. Correlating light scattering with internal cellular structures [J]. *Biomedical Optics Express*, 2012, 3(2): 296-312.
- [14] SU X, GUPTA M, CAPJACK C E, et al. Label-free and noninvasive optical detection of the distribution of nanometer-size mitochondria in single cells[J]. *Journal of Biomedical Optics*, 2011, 16(6): 067003.
- [15] LIU J, XU Y, WANG W, et al. Machine learning of diffraction image patterns for accurate classification of cells modeled with different nuclear sizes[J]. *Journal of Biophotonics*, 2020, 13(9): e202000036.
- [16] SU X, LIU S, QIAO X, et al. Pattern recognition cytometry for label-free cell classification by 2D light scattering measurements[J]. *Optics Express*, 2015, 23(21): 27558-27565.
- [17] CHICEA D. An artificial neural network assisted dynamic light scattering procedure for assessing living cells size in suspension[J]. *Sensors*, 2020, 20(12): 3425.
- [18] 杨斌, 张驰, 平力, 等. 循环流化床颗粒团多参数的光散射测量方法[J]. *仪器仪表学报*, 2021, 42(10): 20-26.
YANG B, ZHANG CH, PING L et al. Light scattering measurement method for multi-parameters of particle clusters in circulating fluidized bed[J]. *Chinese Journal of Scientific Instrument*, 2021, 42(10): 20-26.
- [19] WANG Z D, DA P L, LIU M, et al. Detection of circulating tumor cells in peripheral blood by a novel two-step tumor cell enrichment-identification system that combines the isolation by size of tumor cells technology and the immunofluorescence technique [J]. *Journal of Cancer Science and Clinical Therapeutics*, 2019, 3(4): 186-210.
- [20] WAQAR H, NOUMAN R, DAANIAL K Y. A sequence-based predictor of zika virus proteins developed by integration of PseAAC and statistical moments [J]. *Combinatorial Chemistry & High Throughput Screening*, 2020, 23(8): 797-804.

作者简介



周雨宁, 2019年于东北大学获得学士学位, 现为天津大学硕士研究生, 主要研究方向为光学成像、机器学习。

E-mail: 970906146@qq.com

Zhou Yuning received her B. Sc. degree from Northeastern University in 2019. Now she is a M. Sc. candidate in Tianjin University. Her main research interests include optical imaging and machine learning.



蒋景英(通信作者), 分别于1996年与1999年于天津医科大学获得学士学位和硕士学位, 2003年于天津大学获得博士学位, 现为北京航空航天大学副研究员, 主要研究方向为光学分子影像和生物医学信号和图像处理。

E-mail: jingyingjiang@buaa.edu.cn

Jiang Jingying (Corresponding author) received her B. Sc. degree in 1996 and M. Sc. degree in 1999 from Tianjin Medical University, and Ph. D. degree from Tianjin University in 2003, respectively. Now she is an associate professor in Beihang University. Her main research interests include optical molecular imaging and biomedical signal and image processing.