

基于机器学习的辅助诊断算法在体内 尿路结石成分鉴别中的应用

高亚明¹ 刘兆邦² 陈斌¹ 李铭² 黄来剑^{3*}

¹(扬州大学信息工程学院 江苏 扬州 225127)

²(中国科学院苏州生物医学工程技术研究所 江苏 苏州 215163)

³(温州市人民医院 浙江 温州 325699)

摘要 利用结石 CT 图像结合机器学习算法进行体内尿路结石的成分分析具有一定的研究价值。对二维和三维图像分别进行灰度调整及序列插值以完成预处理操作;利用临床医生已勾勒病灶的标签文件进行图像分割,并对感兴趣区域分别进行二维和三维上的灰度、形状、纹理特征提取;对比不同特征选择算法和分类器,使用 mRMR 算法并结合 RBF_SVM 分类器进行训练。实验结果表明,该算法能有效分析出体内尿路结石中的单纯草酸钙和无水尿酸成分,ACC 和 AUC 分别达到 81.76% 和 89.03%,可以为临床医生诊断提供有效的参考依据。

关键词 结石 CT 图像 机器学习 结石成分 辅助诊断

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2020.12.021

APPLICATION OF MACHINE LEARNING-BASED ASSISTED DIAGNOSTIC ALGORITHMS IN THE IDENTIFICATION OF URINARY CALCULI COMPONENTS IN VIVO

Gao Yaming¹ Liu Zhaobang² Chen Bin¹ Li Ming² Huang Laijian^{3*}

¹(College of Information Engineering, Yangzhou University, Yangzhou 225127, Jiangsu, China)

²(Suzhou Institute of Biomedical Engineering Technology, Chinese Academy of Sciences, Suzhou 215163, Jiangsu, China)

³(Wenzhou People's Hospital, Wenzhou 325699, Zhejiang, China)

Abstract It has certain research value to use CT images of calculi combined with machine learning algorithm for calculi composition analysis in vivo. Two-dimensional and three-dimensional images were adjusted by gray scale and interpolated by sequence to complete the pre-processing operation; the image was segmented by the label files of the lesions outlined by clinicians, and the gray, shape and texture features of the regions of interest were extracted on two-dimensional and three-dimensional respectively; We compared with different feature selection algorithms and classifiers, and mRMR algorithm combined with RBF_SVM classifier was used for training. The experimental results show that the algorithm can effectively analyze the simple calcium oxalate and anhydrous uric acid in urinary calculi, and ACC and ACU reach 81.76% and 89.03% respectively, which can provide an effective reference for clinical diagnosis.

Keywords CT image of calculi Machine learning Stone composition Assistant diagnosis

0 引言

近年来,我国人民的生活质量逐步提高,尿路结石

的发病率也随之递增,南方发病率高于北方,约为 22% ~ 45%,个别省市甚至高达 50% 以上,且 25% 的患者需要住院治疗。显而易见,尿路结石已成为重要的公共卫生问题^[1]。国内大多数研究表明,尿路结石

的形成与很多因素有关,如地理环境、气候、饮食习惯等。本文主要研究体内结石中的单纯性结石,其成分以草酸钙最为多见,占结石的 80% 以上,无水尿酸次之^[2]。目前利用传统的结石成分分析方法对两者进行区分需要借助仪器,步骤复杂且成本较高,无法实现术前体内无创检测。因此,本文通过 CT 图像,结合机器学习算法对两者的结构进行分析鉴别^[3]。

就目前而言,基于机器学习的辅助诊断算法已经从纯粹的理论发展到了临床试验,在诸多方面如乳腺、肺结节等诊断上证实了其可行性与适用性^[4],但其在结石成分鉴别上的研究较少,比如 Perrot 等^[5]利用放射和机器学习能准确鉴别肾结石和静脉增生,却无法有效地在结石间进行鉴别。因此,针对体内单纯性结石特征,本文将辅助诊断算法用于其成分分析中,辅助医生进行诊断。

1 图像预处理

预处理主要运用图像增强技术,传统的图像增强技术大多是基于空间域对图像进行处理,本文利用灰度调整和图像插值方法。

1.1 灰度调整

灰度调整利用灰度直方图得到图像中的像素亮度分布情况,再通过均衡化、规范化处理,调整图像的清晰度。

本文首先将 CT 图像和普通像素值更改为医学专用的 CT 值,再根据结石组织 CT 值范围截取 [100, 1 500]HU 以内的内容进行灰度调整^[6],其中 HU (Hounsfield Unit) 值表示组织对 X 射线的吸收程度。调整效果如图 1 所示。

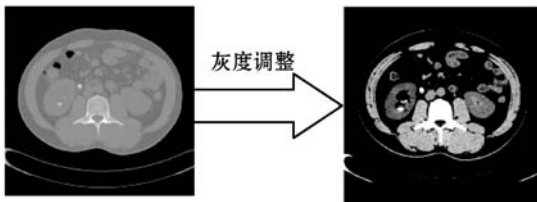


图 1 灰度调整

1.2 图像间插值

图像间插值在每帧图像中增加插值图像以缩小距离,并将插值图像和原始图像共同组成三维数据以提高三维图像的质量^[7]。

本文利用立方插值算法在图像序列间进行超分辨率重建,将 CT 图像的像素间距统一为 $0.7 \text{ mm} \times 0.7 \text{ mm}$,如图 2 所示。

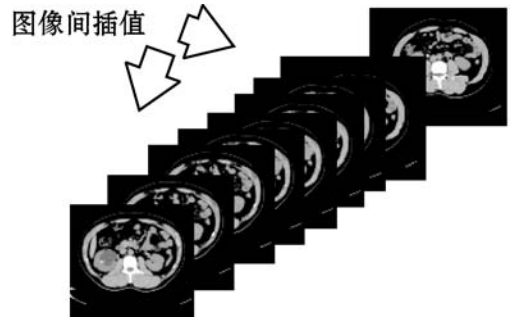


图 2 图像间插值

2 图像分割

当前医学图像分割正从手动或半自动向全自动发展^[8]。本文将临床专家人工分割的结果作为金标准,利用 3D Slicer 软件对病灶组织手动勾勒感兴趣区域 (ROI)^[9]并对其裁剪得到更精准的 ROI 窗口^[10],分割过程如图 3 所示。

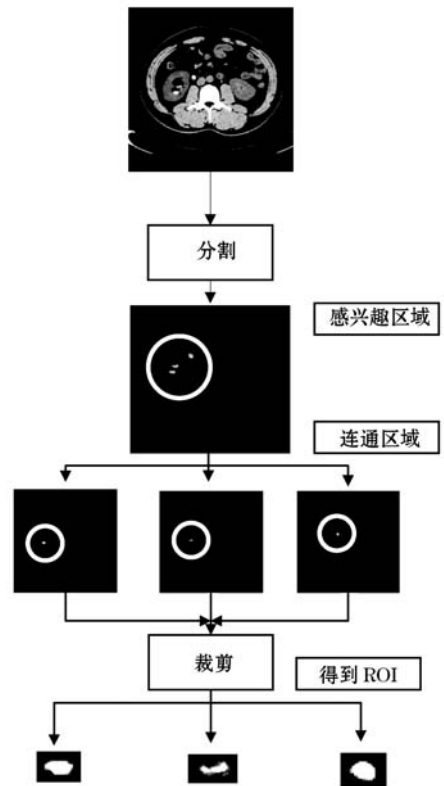


图 3 图像分割

3 特征提取

由于体外的草酸钙结石质硬,呈环形或桑葚形,表面粗糙有刺,而无水尿酸结石质硬且表面光滑,呈圆形或卵圆形^[11],故本文针对体内结石 CT 图像,分别提取在二维和三维上的灰度、形状和纹理特征^[12],并对其进行对比分析。

3.1 灰度特征

医学图像中的灰度直方图考虑形成图像像素的强度,定义如下:由 $f(x,y)$ 表示一幅灰度图像, (x,y) 处的值表示该位置像素,一幅图有 M 个像素,分布在 $0 \sim L-1$ 灰度级,灰度直方图则统计每级灰度像素数目得到的统计图。

$$H(i) = \frac{m_i}{M} \quad i=0,1,\dots,L-1 \quad (1)$$

式中: i 表示灰度级; L 表示灰度级种类数($L \leq 256$); m_i 表示灰度级为 i 的像素个数; M 表示图像总像素个数。

本文利用灰度直方图,得到的灰度特征如下:

- 1) 最大灰度值:分布在 CT 图像中的最大灰度级。
- 2) 均值:图像灰度平均值。

$$\mu = \sum_{i=0}^{L-1} iH(i) \quad (2)$$

- 3) 方差:图像灰度在数值上的分布情况。

$$\sigma^2 = \sum_{i=0}^{L-1} (i - \mu)^2 H(i) \quad (3)$$

- 4) 熵:直方图分布的均匀性。

$$G = \sum_{i=0}^{L-1} H(i) \log_2 H(i) \quad (4)$$

- 5) 倾斜度:直方图分布的不对称程度。

$$\mu_s = \frac{1}{\sigma^2} \sum_{i=0}^{L-1} (i - \mu)^3 H(i) \quad (5)$$

- 6) 峰度:图像的灰度分布在接近均值时的大致状态。

$$\mu_k = \frac{1}{\sigma^4} \sum_{i=0}^{L-1} (i - \mu)^4 H(i) - 3 \quad (6)$$

3.2 形状特征

形状特征为结石分析提供重要的结构信息,分为轮廓特征和区域特征。轮廓特征主要针对结石的外边界,本文得到基于轮廓的特征如下:

- 1) 紧密度:衡量一个形状的紧致程度。

$$Compactness = \frac{A^2}{2\pi \sqrt{i_1^2 + i_2^2}} \quad (7)$$

式中: A 为形状面积; i_1 和 i_2 为二阶矩:

$$i_1 = \sum_I \sum_J i^2 image(i,j)$$

$$i_2 = \sum_I \sum_J j^2 image(i,j)$$

式中: $image(i,j)$ 为 CT 图像; (i,j) 处的值为 CT 图像中该位置的像素值; I 为 CT 图像中所有像素点的横坐标集合; J 为 CT 图像中所有像素点的纵坐标集合。

- 2) 最长径:边界上相距最远的两个点之间的距离。

- 3) 离心率:用焦点间的距离除以长轴的长度。

区域特征则针对整个结石区域,本文得到基于区域的特征包括二维区域中的面积和三维区域中的体积。

3.3 纹理特征

纹理特征反映结石表面的粗糙度、光滑性^[13],提取方法分为统计分析、模型分析、结构分析和频谱分析。

本文采用基于统计的方法,主要包括:

- 1) 灰度共生矩阵(GLCM):统计图像中不同灰度值的像素对出现的情况。

- 2) 灰度游程矩阵(GLRLM):统计图像中具有相同灰度的直线区域(灰度、方向、长度)出现的情况。

- 3) 灰度区域尺寸矩阵(GLSZM):统计图像中具有相同灰度的区域(灰度、大小)出现的情况。

- 4) 邻域灰度差分矩阵(NGTDM):统计图像中相邻区域的灰度差异情况。

- 5) 灰度差分统计(GRAY_DIFF):反映目标场景在连续时间点图像相减所构成的图像特征。

其中灰度共生矩阵(GLCM)最为经典,应用最广,由它提取出如下特征:

- 1) 角二阶矩(energy),反映图像区域的均匀性或平滑性,表示为:

$$w_1 = \sum_i \sum_j [m(i,j)]^2 \quad (8)$$

- 2) 对比度(contrast),表示图像的纹理清晰度,表示为:

$$w_2 = \sum_i \sum_j (i-j)^2 m(i,j) \quad (9)$$

- 3) 相关系数(correlation),反映矩阵行与列的线性相关程度,值越大图像区域灰度分布越均匀,表示为:

$$w_3 = \frac{\sum_i \sum_j ij m(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (10)$$

式中: μ_x 、 μ_y 、 σ_x 、 σ_y 分别定义为:

$$\mu_x = \sum_i i \sum_j m(i,j)$$

$$\mu_y = \sum_j j \sum_i m(i,j)$$

$$\sigma_x^2 = \sum_i (i - \mu_x)^2 \sum_j m(i,j)$$

$$\sigma_y^2 = \sum_j (j - \mu_y)^2 \sum_i m(i,j)$$

- 4) 熵(entrop),度量图像内容的随机性,表示为:

$$w_4 = - \sum_i \sum_j m(i,j) \log_2 [m(i,j)] \quad (11)$$

- 5) 逆差分矩(uniformity),定义为:

$$w_5 = \sum_i \sum_j \frac{m(i,j)}{1 + (i-j)^2} \quad (12)$$

6) 同质化(homogeneity),定义为:

$$w_6 = \sum_i \sum_j \frac{m(i,j)}{1+|i-j|} \quad (13)$$

7) 和平均(sum average),定义为:

$$w_7 = \sum_k \sum_i \sum_j i \cdot im(i,j) \quad (14)$$

式中: $k = i + j$ 。

所提特征如表 1 所示。

表 1 所提特征

特征描述	特征组成	特征维度	特征个数
灰度特征	CTmax	2D&3D	6&6
	CTmean		
	方差		
	熵		
	倾斜度		
	峰度		
形状特征	紧密度	2D&3D	4&4
	面积/体积		
	最长径		
	离心率		
纹理特征	灰度差分统计	2D&3D	55&55
	灰度共生矩阵		
	局部灰度差分矩阵		
	灰度区域尺寸矩阵		
	灰度游程矩阵		

特征提取伪代码如下:

功能:特征提取。

输入:路径字符串 1,路径字符串 2。

输出:2D 灰度特征,2D 形状特征,2D 纹理特征,3D 灰度特征,3D 形状特征,3D 纹理特征。

从磁盘读取 dcm 医学图像数据文件

从磁盘读取 nrrd 医学图像标签文件

for $x = 3$ to $patient_num$

 获取 dcm 文件与 nrrd 文件的数量 img_num

 for $i = 1$ to img_num

 按照 InstanceID 对 dcm 文件进行排序,保证与 nrrd 文件顺序一致

 end

for $i = 1$ to img_num

 查看当前帧是否存在标签

 if $m1 > 0$ then

 对当前帧图像预处理

 for $y = 1$ to num

 对当前帧图像的每个连通区域进行分割,得到 ROI

 对 ROI 进行裁剪得到最大 ROI,即 CT

将所有图像 Im 和标签 mm 分别保存为三维矩阵 O_CT3V 、 $MSK3V$

end

end

end

对 CT 提取二维灰度、形状、纹理特征

将所提特征归类为二维全局特征和内部特征

对 O_CT3V 、 $MSK3V$ 进行裁剪

 计算维度 dim

 if $dim > 1$

 对三维图像预处理

 for $y = 1$ to num

 对每个三维连通区域进行分割,得到 ROI

 遍历得到最大的 ROI,并存储为 $CT3V$

 end

 for $z = 1$ to $size(CT3V, 3)$

 过滤,剔除没有标签的帧图像

 end

 对 $CT3V$ 提取三维灰度、形状、纹理特征

 将所提特征归类为三维全局特征和内部特征

 else

 只有一帧图像,直接将三维当作二维处理,得到特征

 end

end

4 特征选择

特征选择是对高维特征进行筛选以降低特征维度的过程^[14],由于最小冗余最大相关(mRMR)算法综合考虑了相关性和冗余性,故本文使用 mRMR 算法。

mRMR 算法通过计算特征之间和类标签之间的互信息来选出冗余性最小和相关性最大的前 N 个特征,即从最小冗余和最大相关出发,给出一种基于互信息的评价准则。其中:最大相关指的是特征与类标签信息之间的相关度最大;最小冗余指的是特征之间的冗余度最小^[15]。

此外,将 mRMR 算法与相关系数(Spearma)法、Relief 算法、SVM_RFE 算法进行对比分析,从而验证其在本文中的适用性。

5 分类

本文采集的 CT 图像数据是有限的,而支持向量机(SVM)分类器本身能较好地解决小样本、高维度等问题,故本文选择 SVM 分类器,针对有限样本进行训练。

在核函数的选择上,基于高斯径向基核函数的高维映射能力强、样本适用性好、多用于特征数量小(15

维)、样本数量正常(119 例)的情况,故本文使用基于高斯径向基核函数的支持向量机(RBF_SVM)分类器。具体实现则采用第三方 LIBSVM 工具包,利用其内置参数进行调参。

同时,将其与基于 Linear 核函数的 SVM(Linear_SVM)分类器,随机森林(RF)分类器和 Adaboost 分类器进行对比分析,从而验证 RBF_SVM 分类器在本文中的适用性。

6 实验

6.1 实验数据和环境

1) 实验所用数据:

从温州市人民医院采集结石 CT 图像 DICOM 数据集,其中:草酸钙患者 59 例;无水尿酸患者 60 例;层内像素分辨率为 512 × 512,层距为 5 mm。

2) 实验所用机器环境:

操作系统:Windows 10;CPU: Intel(R) Core(TM) i5 - 7500 CPU @ 3.40 GHz(3 401 MHz);内存:16 GB;编程软件:MATLAB R2018a。

6.2 评估方法

1) 分层 10 次 10 折交叉验证法。分层 10 折交叉验证首先将数据集分成 10 份,每份中类别之间的比例与整个数据集中的比例相同,轮番将其中 9 份作为训练集,1 份作为测试集进行实验,每轮得到一个精度,最后对 10 轮结果取平均作为最后的模型精度。模型精度估计需要进行多次 10 折交叉验证,本文进行 10 次 10 折交叉验证,对得到的 10 个精度(ACC)求均值。

2) 调参。参数设定直接影响算法性能,本文利用交叉验证法,结合 ROC 曲线下的面积(AUC)对各分类器参数设置步长,进行调参,以选取最优参数,优化分类模型。

3) 性能度量。混淆矩阵是一个误差矩阵,由它可得许多度量指标,如表 2 所示。

表 2 混淆矩阵

真实情况	预测结果	
	正例	反例
正例	TP(真正例)	FN(假反例)
反例	FP(假正例)	TN(真反例)

$$\text{精度 } ACC = \frac{\text{分类正确的样本数}}{\text{样本总数}} \quad (15)$$

$$\text{查准率 } PRE = \frac{TP}{TP + FP} \quad (16)$$

$$\text{查全率 } RECALL = \frac{TP}{TP + FN} \quad (17)$$

$$F \text{ 度量} = \frac{2 \times TP}{\text{样例总数} + TP - TN} \quad (18)$$

$$\text{灵敏度 } SEN = \frac{TP}{TP + FN} \quad (19)$$

$$\text{特异性 } SPE = \frac{TN}{TN + FP} \quad (20)$$

ROC 曲线的横轴为假正例率:

$$FPR = \frac{FP}{TN + FP} \quad (21)$$

ROC 曲线的纵轴为真正例率:

$$TPR = \frac{TP}{TP + FN} \quad (22)$$

根据分类器的预测概率迭代更新阈值,得到多个混淆矩阵对应的坐标点,依据坐标点画出 ROC 曲线进而计算 AUC 值。

6.3 实验步骤与结果

如图 4 所示,首先对二维图像进行灰度调整,对三维图像序列进行插值,完成超分辨率重建,统一分辨率;其次,利用已标好病灶的 NRRD 标签文件进行图像分割,并裁剪得到更精准的 ROI 窗口;然后,再对 ROI 分别提取二维和三维上的灰度、形状、纹理特征,并通过特征选择算法进行筛选;最后将筛选得到的特征放入分类器进行模型训练,通过 ACC、AUC 等指标对比分析分类效果。

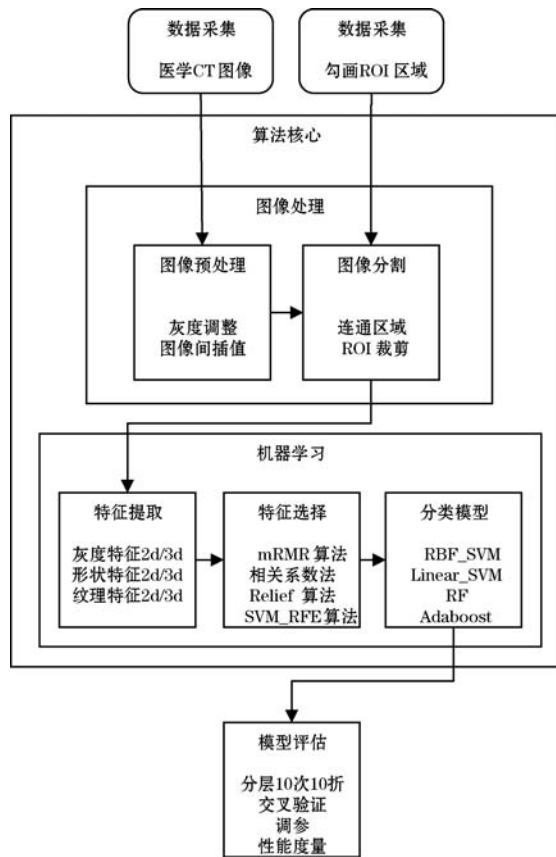


图 4 实验步骤

1) 特征提取算法对比。如表 3 和表 4 所示,各领域信息互补,综合训练模型的鉴别性能,在统一使用 mRMR 算法与 RBF_SVM 分类器的前提下,纹理特征比灰度和形状特征鉴别能力更强,2D 特征略优于 3D 特征,而提取所有特征时,ACC 值和 AUC 值最高,即最能反映病灶组织空间分布的异质性。

表 3 特征提取平均 ACC 比较 %

维度	灰度	形状	纹理	所有
2D	76.64	66.89	76.30	80.92
3D	78.99	54.87	74.53	79.49
所有	77.14	63.95	78.73	81.76

表 4 特征提取平均 AUC 比较 %

维度	灰度	形状	纹理	所有
2D	84.28	65.66	82.28	87.58
3D	86.84	54.82	79.01	86.69
所有	83.57	65.97	84.06	89.03

2) 特征选择算法对比。对各特征选择算法进行对比分析,由表 5 可知,在统一提取所有特征与使用 RBF_SVM 分类器的情况下,mRMR 综合考虑了特征之间的冗余性和特征类标之间的相关性,故其选择效果最好,验证了 mRMR 在本文中的适用性。

表 5 特征选择对比 %

算法	平均 ACC	平均 AUC
mRMR	81.76	89.03
Spearma	81.69	88.20
SVM_RFE	81.62	88.44
Relief	81.18	88.27
不选择	80.34	87.35

3) 分类器对比。由表 6 可知,在统一提取所有特征与使用 mRMR 的情况下,RBF_SVM 因其在小样本上的适用性较强而分类效果最好,验证 RBF_SVM 在本文中的适用性。

表 6 分类器对比 %

分类器	平均 ACC	平均 AUC
RBF_SVM	81.76	89.03
Linear_SVM	80.85	88.97
RF	80.42	88.64
Adaboost	77.05	87.45

4) 最终辅助诊断算法评估。对结石 CT 图像预处理及分割后提取所有特征,共 130 维,再利用 mRMR 算法筛选,筛选特征数对比如图 5 所示。

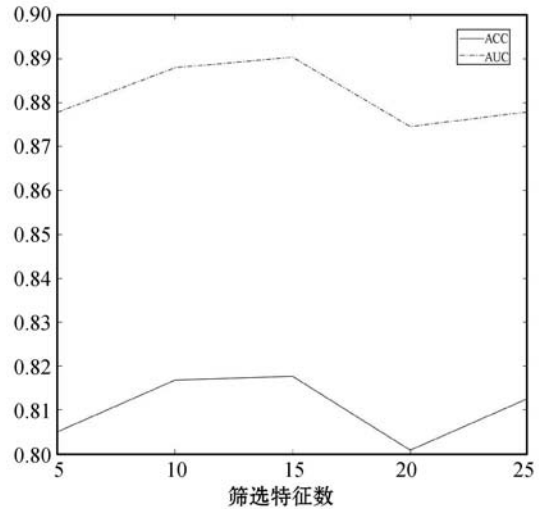


图 5 筛选特征数对比

由图 5 可知,选择前 15 个特征分类效果最好,所选特征如表 7 所示。

表 7 所选特征

序号	下标	特征	维度
1	2	CTmean	2D
2	35	粗糙度	2D
3	46	高灰度区域因子	2D
4	44	区域百分比	2D
5	66	CTmax	3D
6	122	游程百分比	3D
7	67	CTmean	3D
8	31	聚块阴影度	2D
9	60	短游程低灰度因子	2D
10	78	能量	3D
11	119	长游程因子	3D
12	26	平方方差和	2D
13	73	体积	3D
14	1	CTmax	2D
15	114	大区域低灰度因子	3D

将所选特征放入 RBF_SVM 分类器,得到本文所述的完整辅助诊断算法模型,综合各项指标对结果进行评估,评估结果如表 8 所示。

表 8 辅助诊断算法的各项指标 %

序号	ACC	PRE	REC ALL	F 度量	SEN	SPE	AUC
1	79.83	82.14	76.67	79.31	76.67	83.05	87.97
2	84.87	87.50	81.67	84.48	81.67	88.14	90.45
3	80.67	82.46	78.33	80.34	78.33	83.05	89.46
4	84.87	85.00	85.00	85.00	85.00	84.75	89.52
5	81.51	85.19	76.67	80.70	76.67	86.44	89.38
6	81.51	82.76	80.00	81.36	80.00	83.05	89.89
7	79.83	83.33	75.00	78.95	75.00	84.74	87.43
8	82.35	83.05	81.67	82.35	81.67	83.05	88.53
9	81.51	82.76	80.00	81.36	80.00	83.05	89.10
10	80.67	83.64	76.67	80.00	76.67	84.74	88.62
平均	81.76	83.78	79.17	81.38	79.17	84.40	89.03

如图 6 和图 7 所示,10 次结果的 ACC 值约为 0.81,说明分类正确的样本数较多,而 ROC 曲线对应的 AUC 值在 0.89 左右,说明分类器性能较优,所以,基于该辅助诊断算法可以实现对草酸钙和无水尿酸结石的准确分类。

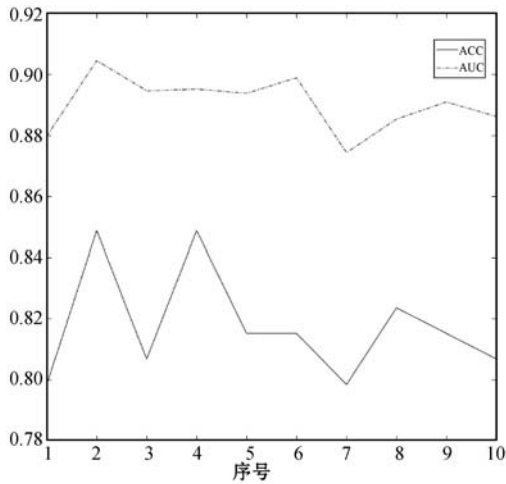


图 6 辅助诊断算法的 ACC, AUC 指标

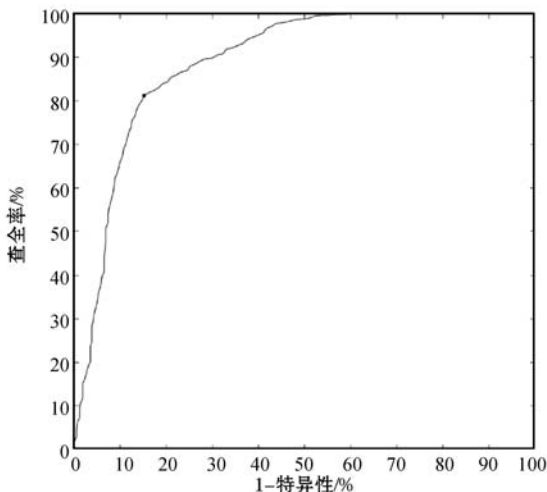


图 7 辅助诊断算法的 ROC 曲线

7 结 语

本文结合临床医学,利用 CT 图像进行定量分析,并对比机器学习特征提取、选择及分类算法的实验结果,提取结石在二维和三维上的灰度、形状、纹理特征。最终使用 mRMR 特征选择算法及 RBF_SVM 分类器,实现患者体内草酸钙和无水尿酸结石的准确分类,ACC 和 AUC 分别达到 81.76% 和 89.03%,比现有结石分析方法更为快速简便,真正实现体内无创,且符合临床统计学结果,为临床医生诊断提供更为有效的参考依据。下一步将优化图像处理和机器学习算法,得到更高的准确率。

参 考 文 献

- [1] 赵晓俊,浦金贤,平季根,等.螺旋 CT 判断上尿路结石化学成分的研究[J].临床医药文献电子杂志,2016,3(51):10089-10090.
- [2] 朱华平,孙琼,杨艳玲.曲靖市 2023 例尿路结石成分分析[J].中外医学研究,2019,17(3):173-174.
- [3] 陈泉桦,黎军强.双源 CT 双能量成像识别尿路结石成分的临床应用研究[J].中国 CT 和 MRI 杂志,2018,16(7):133-135,139.
- [4] 梁长虹.加强机器学习在医学影像中的研究和应用[J].国际医学放射学杂志,2019,42(1):1-2.
- [5] Perrot T D, Hofmeister J, Burgermeister S, et al. Differentiating kidney stones from phleboliths in unenhanced low-dose computed tomography using radiomics and machine learning[J]. European Radiology, 2019, 29:4776-4782.
- [6] 范钰,刘振华,魏强,等.上尿路结石 CT 值预测结石成分和体外冲击波碎石术治疗的可行性分析[J].四川大学学报(医学版),2017,48(5):788-791.
- [7] 王兵,樊星,杨颖,等. CT 图像肺结节的三维超分辨率重建及显示[J].生物医学工程学杂志,2015,32(4):788-794.
- [8] Wei B, Zheng Y, Zhang K. An fMCMC medical image segmentation algorithm[J]. Journal of Medical Imaging and Health Informatics, 2017, 7(5):1057-1062.
- [9] Kalshetti P, Bundele M, Rahangdale P, et al. An interactive medical image segmentation framework using iterative refinement[J]. Computers in Biology and Medicine, 2017, 83(1):22-33.
- [10] Lu Y W, Jiang J G, Qi M B, et al. Segmentation method for medical image based on improved grabcut[J]. International Journal of Imaging Systems and Technology, 2017, 27(4):383-390.

可以看出,应用 CAP-AFSA 调度 RG1、RG2、RG4 的迭代曲线分别在节能效率达到 60.7%、54.3%、62.2% 的附近时寻优过程暂时停滞,而后均跳出该点并进一步迭代寻优。因此,可以认为执行 RG1、RG2、RG4 的节能效率局部最优值分别位于 60.7%、54.3%、62.2% 附近,这恰好是应用 CAPSER 执行 RG1、RG2、RG4 所得的节能效率。可见,应用 CAP-AFSA 寻优时有效地跳出了局部最优,而应用 CAPSER 寻优时陷入了局部最优。上述仿真结果表明,与 GA 相比,AFSA 的全局寻优能力更强,从而应用 CAP-AFSA 能较好地跳出局部最优。

4 结 语

本文针对基于 GA 的调度策略 HGLS + PDP - SPM、CASPER 存在的易陷入局部最优等突出问题,基于有约束的 AFSA,研究多核系统的实时节能调度问题。通过定义并基于任务节点的高度值确定任务的执行顺序;对任务分配和电压调节进行有机的整合编码;对人工鱼位置向量进行取整和越界处理;基于可行性规则处理截止期约束以保证模拟行为的可行性,进而限制寻优过程等一系列过程,建立有效的问题模型,并应用 AFSA 进行迭代寻优。仿真结果表明,与 HGLS + PDP - SPM、CASPER 相比,应用 CAP-AFSA 执行任务图可以获得更好的节能效果,其总体性能有较显著的改善,因而不失为一种有效的、更具应用性的实时节能调度策略。下一步工作将集中于对更为复杂的多核系统的实时节能调度问题模型进行研究。

参 考 文 献

[1] 张冬松,赵志峰,林志辉,等. 基于多核处理器系统多维限制的节能实时调度[J]. 计算机系统应用,2015,24(10): 204 - 211.

[2] Xian C J, Lu Y H, Li Z Y. Energy-aware scheduling for real-time multi-processor systems with uncertain task execution time[C]//2007 44th ACM/IEEE Design Automation Conference. IEEE,2007:664 - 669.

[3] Pagani S, Chen J J. Energy efficiency analysis for the single frequency approximation(SFA) scheme[J]. ACM Transactions on Embedded Computing Systems,2014,13(5S): 158.

[4] 王宇. 基于 DVS 的多核周期任务节能调度策略研究[D]. 武汉:武汉理工大学,2013.

[5] Huang X, Li K L, Li R F. An energy efficient scheduling base on dynamic voltage and frequency scaling for multi-core embedded real-time system[C]//Proceedings of the 9th International Conference on Algorithms and Architectures for

Parallel Processing. Springer,2009: 137 - 145.

- [6] Kianzad V, Bhattacharyya S S, Qu G. CASPER: an integrated energy driven approach for task graph scheduling on distributed embedded systems[C]//16th IEEE International Conference on Application-Specific Systems, Architectures, and Processors. IEEE,2005:191 - 197.
- [7] 李晓磊,邵之江,钱积新. 一种基于动物自治体的寻优模式:鱼群算法[J]. 系统工程理论与实践,2002,22(11): 32 - 38.
- [8] 许少尉,吕浩. 多核处理器系统节能调度技术研究[J]. 航空计算技术,2018,48(1):98 - 101.
- [9] Dick R P, Rhodes D L, Wolf W. TGFF: Task graphs for free[C]//6th International Workshop on Hardware/Software Codesign. IEEE,1998:97 - 101.
- [10] 葛宇翔. 基于多核环境的节能调度策略研究[D]. 武汉:武汉理工大学,2017.
- [11] 赵阳. 多核系统中基于 DVS 的实时节能调度方法研究[D]. 株洲:湖南工业大学,2013.
- [12] Al-Maasarani A. Priority-based scheduling and evaluation of precedence graphs with communication times[D]. Saudi Arabia: University of Petroleum and Minerals,1993.
- [13] Al-Mouhamed M A. Lower bound on the number of processors and time for scheduling precedence graphs with communication costs[J]. IEEE Transactions on Software Engineering,1990,16(12):1390 - 1401.
- [14] McCreary C L, Khan A A, Thompson J J, et al. A comparison of heuristics for scheduling DAGS on multiprocessors [C]//8th International Parallel Processing Symposium. IEEE,1994:446 - 451.
- [15] Wu M Y, Gajski D D. Hypertool: A programming aid for message-passing systems[J]. IEEE Transactions on Parallel and Distributed Systems,1990,1(3):330 - 343.

(上接第 139 页)

- [11] 董浩,王琦,彭泳涵,等. 泌尿系上尿路结石成分的单中心研究[J]. 第二军医大学学报,2018,39(4):450 - 454.
- [12] 刘晓虹,朱玉全,刘哲,等. 基于改进多尺度 LBP 算法的肝脏 CT 图像特征提取方法[J]. 计算机科学,2019,46(3): 125 - 130.
- [13] Sun Q,Zhao D,Cheng S, et al. A feature extraction method for adaptive DBS using an improved EMD [J]. International Journal of Neuroscience,2018,128(10):975 - 986.
- [14] 张飞飞,周涛,陆惠玲,等. 基于贝叶斯粗糙集的肺部肿瘤 CT 图像高维特征选择算法[J]. 生物医学工程研究,2018,37(4):404 - 409.
- [15] Zerhari B, Lehen A A, Mouline S. A new horizontal distributed feature selection approach [J]. Cybernetics and Information Technologies,2018,18(4):15 - 28.