

通过 PET-CT 图像纹理特征预测软组织肉瘤转移性

申俊丽,余堃

(河南师范大学 计算机与信息工程学院,河南 新乡 453007)

摘要:提出了一种针对软组织肉瘤转移性预测的辅助诊断方法,该方法通过对患者的 FDG-PET 和 CT 诊断图像进行纹理特征分析,共提取了 105 个特征,其中包括灰度共生矩阵的 24 个特征和其他 81 个灰度等级的特征,分别利用支持向量机、K 近邻和随机森林等机器学习算法建立预测模型,并采用网格搜索法对其参数进行优化.最后使用留一交叉验证法对各模型进行验证.通过评估各模型性能,选择支持向量机作为最终预测模型,得到了 80% 的平均精确度.此外,该模型的敏感度达到 81%,特异性达到 79%,表明该模型预测结果具有一定的可靠性,可以对 STS 进行辅助诊断并通过更好的适应性治疗来改善患者的预后.

关键词:软组织肉瘤;纹理特征;机器学习;转移性预测

中图分类号:TP181;TP391

文献标志码:A

随着人类生存环境的不断恶化,癌症的发病率也在逐年攀升.软组织肉瘤(soft tissue sarcoma, STS)因其倾向于在早期阶段发生转移的特性而受到广泛关注.STS 是一组具有侵犯性的软组织病变,不同类型的 STS 出现在身体的不同部位,如神经、肌肉和脂肪,四肢是最常见的起源部位.根据美国癌症协会统计,大约有 50 多种子类型的肉瘤.报告还称,2019 年将会诊断出约 12 750 名新 STS 患者,预计约有 5 270 名患者死于 STS^[1].

通常情况下,通过不同形式的治疗会对四肢 STS 有良好的局部控制.然而,大约 25% 的 STS 患者会发生远处转移^[2],特别是高级别的肿瘤,其转移率上升至约 50%^[3].在转移性肉瘤中,预期平均生存期约为 1 年,2 年生存率为 20%^[4].文献[5]在研究中发现所有发生转移后的患者存活中位数为(20±3.7)月,2 年和 3 年生存率分别为(47.1±4.8)%和(34.4±5.2)%.因此,对 STS 进行早期评估,能够改善患者预后并且提高存活率.

本文利用 PET/CT 成像技术对治疗前图像进行分析,同时结合机器学习建立预测肿瘤是否会发生转移(复发)的模型.首先根据患者临床 PET/CT 图像数据,分析其纹理特征变化,对图像数据进行量化.然后分别采用支持向量机(SVM)、K 近邻(KNN)和随机森林(RF)3 种机器学习算法对量化过的数据进行分类并建立预测模型,通过模型的性能对比选择 SVM 建立的模型为最终预测模型,得到了令人满意的结果.

1 数据集介绍

1.1 FDG-PET/CT 图像

所有的恶性肿瘤都会出现葡萄糖代谢增加的特征,依据这一性质可以使用葡萄糖的类似物 18F-氟脱氧葡萄糖(18F-FDG)对肿瘤进行评估^[6].18F-FDG 是一种含有放射性物质的葡萄糖溶剂,利用正电子发射断层扫描/计算机断层扫描(PET/CT)设备,可以获得 PET 和 CT 的序列图.PET 图像的作用是显示主要部位的代谢信息,CT 图像可以提供解剖结构和密度信息.在 STS 诊断中,FDG-PET 起到非常关键的作用^[7].

收稿日期:2019-07-07;修回日期:2020-02-20.

基金项目:国家自然科学基金青年基金(11601130)

作者简介(通信作者):申俊丽(1979—),女,河南淮阳人,河南师范大学副教授,博士,研究方向为医学图像处理与人工智能,E-mail:zuoyawen1215@126.com.

1.2 数据集

本研究中数据集是通过检索公开可用的癌症成像档案(TCIA)数据库而得到.数据由51名患有经人体组织学证实的四肢 STS 患者的 FDG-PET/CT 和解剖磁共振(MRI)成像数据组成^[8].所有患者在2004年11月至2011年11月期间进行了治疗前 FDG-PET/CT 和 MRI 扫描.表1提供了患者转移具体部位的统计数据.在这项研究中,仅分析了51名患者的数据集,每名患者的 PET 和 CT 图像序列包括约500个格式为 DICOM 的图像,其中,每张 CT 图像的尺寸为 512×512 像素,每张 FDG-PET 图像的尺寸为 128×128 像素.

表1 患者转移部位统计

Tab. 1 Statistics of patients' metastatic sites

转移部位	患者数量/%
肺部	19/37.26
手臂	1/1.96
骨骼	2/3.92
脊柱	1/1.96
腹部	1/1.96
局部复发	3/5.88
未发生转移(复发)	24/47.06

2 放射学特征及提取流程

2.1 放射学特征

本次实验医疗图像采用了来自7个类别的105个放射学特征.具体包括:一阶统计学特征,3D形状特征,灰度共生矩阵(Gray Level Cooccurrence Matrix, GLCM),灰度区域矩阵(Gray Level Size Zone Matrix, GLSZM),灰度行程矩阵(Gray Level Run Length Matrix, GLRLM),邻域灰度差矩阵(Neighbouring Gray Tone Difference Matrix, NGTDM)和灰度依赖矩阵(Gray Level Dependence Matrix, GLDM).

(1)一阶统计特征.一阶统计特征不考虑相邻体素之间的关系,而是分析 ROI(region of interest)内的体素强度,构造直方图.本文采用了包括能量(Energy)、熵(Entropy)、均值(Mean)和峰度(Kurtosis)等在内的16个参数来描述基于直方图的特征.

(2)3D形状特征.3D形状特征描绘 ROI 区域的形状和大小,通常使用体积,表面积等参数来描述 3D 形状特征.这里使用了包括网格体积(Mesh Volume)、表面积(Surface Area)和体素体积(Voxel Volume)等在内的14个参数来描述形状特征.

(3)灰度共生矩阵(Gray Level Cooccurrence Matrix, GLCM).GLCM 是 20 世纪 70 年代由文献[9]首次提出,是从图像中提取纹理信息的统计工具.通过对 ROI 区域内的体素对出现的频次统计得到 GLCM,反应图像局部区域或整体区域的某像素与相邻像素或一定距离内的像素的灰度关系.这里使用了包括自相关(Autocorrelation)、逆差(Id)、逆差矩(Idm)和对比度(Contrast)等24个参数来描述该特征.

(4)灰度区域矩阵(Gray Level Size Zone Matrix, GLSZM).GLSZM 统计了 ROI 区域内 n 个连接的邻域内具有相同强度的特定大小区域的数量^[10].这里使用了包括灰度不均匀性(Gray Level NonUniformity)、区域百分比(Zone Percentage)和区域熵(Zone Entropy)等16个参数来描述该特征.

(5)灰度行程矩阵(Gray Level Run Length Matrix, GLRLM).GLRLM 是用于提取图像中较高阶(两个以上像素之间的空间关系)统计特征的矩阵,表示一组具有相同灰度值的连续共线图像点,行程的长度是图片中运行点的数量^[11].这里使用了包括长行程重点(Long Run Emphasis)、短行程重点(Short Run Emphasis)和行程熵(Run Entropy)等16个参数来描述该特征.

(6)邻域灰度差矩阵(Neighbouring Gray Tone Difference Matrix, NGTDM).NGTDM 刻画了中心体素与其周围像素点的差异^[10],通过计算中心体素与周围体素平均值的差而得到.通常采用5个参数对其进行描述:繁忙度(Busyness)、粗糙度(Coarseness)、复杂度(Complexity)、对比度(Contrast)和强度(Strength).

(7)灰度依赖矩阵(Gray Level Dependence Matrix, GLDM).GLDM 描绘了具有特定像素间距离和方向的灰度对,并使用其空间分布特征来表征图像区域中的纹理^[12].这里使用了包括依赖熵(Dependence Entropy)、依赖方差(Dependence Variance)和大依赖重点(Large Dependence Emphasis)等14个参数来描述该特征.

上述特征的描述以及详细的计算方法可以参考文献[13]提出的生物标志物标准化倡议(IBSI).

2.2 特征提取流程

通常情况下,放射学特征提取需要以下 3 个步骤(如图 1):(1)从专业成像设备得到图像,获得的图像一般有 CT, MRI, X 光等能提供清晰人体解剖结构的图像;(2)绘制病灶区域(ROI),该过程通常会由人工完成,需要专业的放射学医生通过手动的方式将病灶区域(ROI)绘制在图像上;(3)计算特征,分割 ROI 内的图像,通过一定的标准计算 ROI 内特征。



图1 放射学特征提取的一般步骤

Fig.1 General steps of radiological feature extraction

采用手工绘制病灶区域的方法需要专业医师执行,而且当肿瘤存在模糊边界时,同一病灶区域会因不同医师和研究员的主观差异而难以绘制出相同或相似的区域.由于人体组织结构的复杂性,目前为止还没有一种使分割区域重复性很高并且适用于所有医学图像的自动分割算法.因此,本文希望借助能够显示病灶部位信息的医学图像来自动确定 CT 图像的 ROI 区域.PET 图像是一种肿瘤诊断、判断预后和反应监测的重要工具,在临床方面的作用逐渐得到了重视,它可以在图像中显示出肿瘤的具体位置、形状和大小等信息(如图 2).患者的病灶部位在左侧臀部,从图 2 中可以清楚地看出病灶部位。

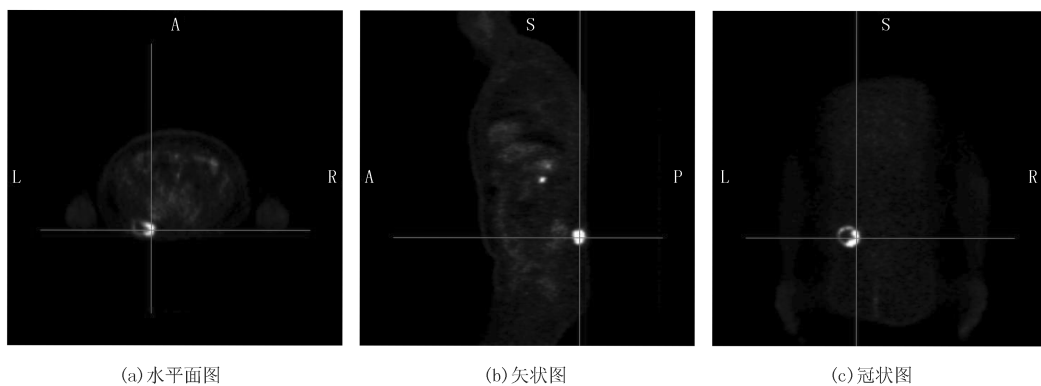


图2 某位患者PET图像

Fig.2 PET image of a patient

FDG-PET 图像和 CT 图像各有优势,如果把这两种图像特征融合起来,能够解决病灶区需要专业医生手动绘制的问题.于是本文提出了如图 3 所示的流程来进行特征提取. FDG-PET 是一种功能信息成像,其空间分辨率为 128×128 像素,而 CT 是一种解剖结构图像,其空间分辨率为 512×512 像素,使得图像不能直接采用 FDG-PET 图像与 CT 图像融合.为了解决这个问题,采用了仿射变换对 FDG-PET 图像进行放大,并且利用线性差值的方法将图像重采样为 512×512 像素.经过处理后, PET 图像的大小与切片数量与 CT 图像的一致。

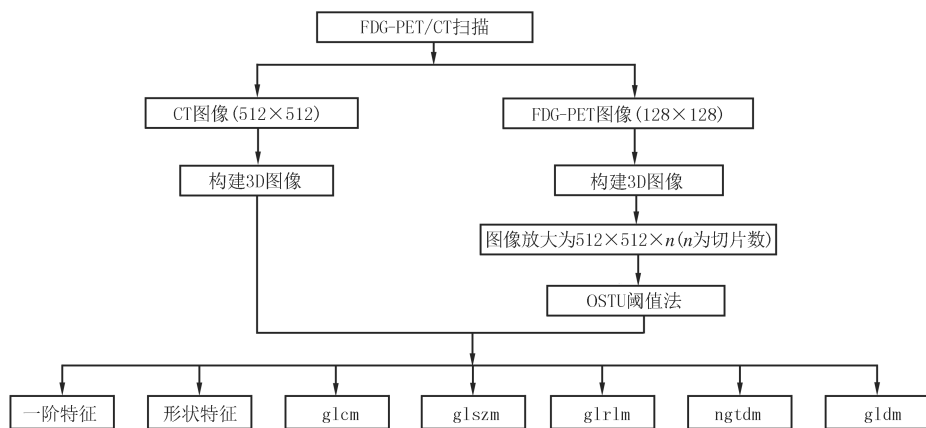


图3 结合PET图像对CT图像提取特征

Fig.3 Feature extraction of CT image combined with PET image

大津算法(OSTU)^[14]是日本学者大津于 1979 年提出的一种非参数无监督的图像分割阈值自动选取方

法.对于放大后的 FDG-PET 图像,采用大津阈值化算法将病灶区域(ROI)分割出来,然后与 CT 图像融合得到病灶区域内的 CT 图像,最后计算图像的放射学特征.

3 预测模型和结果评估

3.1 预测模型

本次实验采用支持向量机(SVM)、K 近邻(KNN)和随机森林(RF)对所提取的特征建立模型进行分类和预测,并通过对比模型性能选择最终的预测模型.

(1)支持向量机(SVM).SVM 是由文献[15]提出的一种基于统计学习理论发展起来的机器学习方法.其主要用于处理小样本数据问题,广泛应用于许多领域,如回归估计、环境光照学习、图像识别和生物信息学分析等,并且在医疗图像分类预测方面也有其独特的优势^[16-19],它可以用如下数学模型表示:

$$\begin{cases} \max_{\lambda} & -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j K(x_i, x_j) + \sum_{i=1}^N \lambda_i, \\ \text{s.t.} & \sum_{i=1}^N \lambda_i y_i = 0, 0 \leq \lambda_i \leq C, i = 1, 2, \dots, N, \end{cases}$$

其中 λ 是拉格朗日乘数, $K(\cdot, \cdot)$ 是核函数, x_i 是第 i 个样本的特征向量, y_i 是 x_i 的对应标签, C 是软边距参数, N 是样本数量.

(2)K 近邻(KNN).K 近邻是一种简单易于实现的机器学习算法,通过根据待分类样本最近的 K 个样本的标签来确定该样本最终的分类标签,是一种常见的分类算法.

(3)随机森林(RF).RF 是一种集成学习算法,使用有放回的抽样方法产生若干组数据集,对每组数据集使用决策树建立模型产生若干决策树,最终的分类结果由这若干决策树的结果投票决定.它具有一定的抗过拟合能力,对数据集的适应能力较强,能够并行处理等优点.

本次实验建立模型的流程具体如图 4 所示:首先,输入 51 名患者的放射学特征,其中每名患者的特征有 105 个;其次,对这些特征进行标准化处理,以提高训练速度和精度;然后,分别运用 3 种机器学习算法建立模型,并对这几种模型使用交叉验证进行评估;最后根据评估结果选择最优的模型作为最终的预测模型.

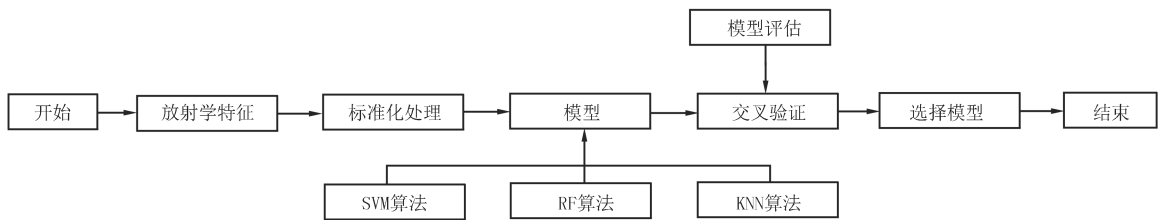


图4 模型建立及选取流程

Fig.4 Model establishment and selection process

一般而言,在模型训练过程中需要对建立的模型进行精度(ACC)评估,这里使用留一交叉验证法对模型进行评估.本次实验依次对提取到的特征分别建立 SVM, KNN 和 RF 预测模型,然后使用网格搜索法对模型进行不断的改进使模型达到理想状态.

通过对模型的不断调整,得到的最终留一法交叉验证的结果如表 2 所示.可以看出,SVM 模型的预测精确度要比 KNN 模型和 RF 模型都要好.

3.2 结果评估

为了进一步量化所提出方法的分类性能,除了衡量精度,还引入了 4 个分类性能指标,包括灵敏度

表 2 留一法交叉验证精度

Tab. 2 Leave-one-out cross-validation accuracy	
算法	精度
SVM	0.80
KNN	0.67
RF	0.65

(SEN),特异性(SPE),马修斯相关系数(MCC)和接收器操作特征(ROC)曲线下面积(AUC)用于本研究。SEN, SPE 和 MCC 定义如下:

$$SEN = \frac{TP}{TP+FN}, SPE = \frac{TN}{TN+FP}, MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}, \text{ 其中}$$

TP(True Positive)表示通过模型正确预测 STS 发生转移的数量;TN(True Negative)表示通过模型正确预测 STS 未发生转移的数量;FP(False Positive)表示通过模型错误预测 STS 发生转移的数量;FN(False Negative)表示通过模型错误预测 STS 未发生转移的数量。

在这几个指标中,AUC 用于量化分类方法的整体分类性能。通常认为,AUC 值越大,分类方法的性能越好。MCC 本质上是一个描述实际分类与预测分类之间的相关系数,它的取值范围为 $[-1, 1]$,取值为 1 时表示对受试对象的完美预测,取值为 0 时表示预测的结果还不如随机预测的结果,-1 表示预测分类和实际分类完全不一致。

通过对上述指标进行计算,得到了表 3。从表 3 中可以看出,使用本文提出的方法建立预测模型相较于仅使用 PET 图像特征或 CT 图像特征建立的模型,其精确度和性能都有明显的优势。

表 3 不同模型性能评估

Tab. 3 Performance evaluation of different models

模型	SEN	SPE	MCC	AUC
SVM	0.81	0.79	0.61	0.65
RF	0.63	0.71	0.34	0.59
KNN	0.59	0.70	0.30	0.59

4 结束语

本文提出了一种结合 PET 图像对 CT 图像进行特征提取的方法,并对使用该方法提取的特征建立 SVM 预测模型。通过精度、灵敏度和特异性等指标的对比,该方法所建立的模型比使用 KNN 和 RF 建立的模型其性能更优。相较于文献[20]提出的通过分析核磁共振图像(MRI)的方法,本方法分类精确度略低(他们的方法精确度为 86.11%),但是不需要放射肿瘤学专家手动地为每个患者定义 3D 肿瘤区域的轮廓,使得该方法能够减少对人工的依赖,预测更有效率。当然本研究还有一定的局限性,首先由于本研究中实验对象的数量太小,不能证明所提出的方法在不同的数据集中具有良好的泛化能力;其次,该实验仅对 STS 发生转移(复发)进行了预测,没有对转移部位进行详细预测;最后,由于使用阈值化算法提取肿瘤区域存在一定的噪声,使得提取的 ROI 区域可能包含正常部位,影响模型预测性能。因此,将尽可能在未来的研究中改善这些局限性,尝试将该方法应用到其他部位肿瘤的预测中。

参 考 文 献

- [1] AMERICAN CANCER SOCIETY. Cancer Facts & Figures[EB/OL].[2019-07-03].<https://w-ww.cancer.org/cancer/soft-tissue-sarcoma/about/key-statistics.html>.
- [2] BILLINGSLEY K G, LEWIS J J, LEUNG D H Y, et al. Multifactorial analysis of the survival of patients with distant metastasis arising from primary extremity sarcoma[J]. Cancer, 1999, 85(2): 389-395.
- [3] BRENNAN M F. Soft tissue sarcoma: Advances in understanding and management[J]. The Surgeon, 2005, 3(3): 216-223.
- [4] BESIROGLU M, DANE F, et al. Systemic chemotherapy of advanced soft tissue sarcomas[J]. Journal of Oncological Sciences, 2017, 3(2): 66-70.
- [5] KANG S, KIM H S, KIM S J, et al. Post-metastasis survival in extremity soft tissue sarcoma: A recursive partitioning analysis of prognostic factors[J]. European Journal of Cancer, 2014, 50(9): 1649-1656.
- [6] AHUJA V, COLEMAN R E, HERNDON J, et al. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with nonsmall cell lung carcinoma[J]. Cancer, 1998, 83(5): 918-924.
- [7] KUBO T, FURUTA T, JOHAN M P, et al. Prognostic significance of 18F-FDG PET at diagnosis in patients with soft tissue sarcoma and bone sarcoma; systematic review and meta-analysis[J]. European Journal of Cancer, 2016, 58: 104-111.

- [8] VALLIÈRES M, FREEMAN C R, SKAMENE S R, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities[J]. *Physics in Medicine & Biology*, 2015, 60(14): 5471-5496.
- [9] HARALICK R M, SHANMUGAM K. Textural features for image classification[J]. *IEEE Trans Syst Man Cybern*, 1973, 3(6): 610-621.
- [10] LYU W, YUAN Q, WANG Q, et al. Robustness versus disease differentiation when varying parameter settings in radiomics features: application to nasopharyngeal PET/CT[J]. *European Radiology*, 2018, 28(8): 3245-3254.
- [11] GALLOWAY M M. Texture analysis using gray level run lengths[J]. *Computer Graphics & Image Processing*, 1975, 4(2): 172-179.
- [12] SVOLOS A E, TODD-POKROPEK A. Time and space results of dynamic texture feature extraction in MR and CT image analysis[M]. Piscataway: IEEE Press, 1998.
- [13] ZWANENBURG A, LEGER S, VALLIÈRES, et al. Image biomarker standardisation initiative-feature definitions[EB/OL]. [2019-07-03]. https://www.researchgate.net/publication/311805734_Image_biomarker_standardisation_initiative_-_feature_definitions.
- [14] OTSU N. A threshold selection method from gray-level histograms[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66.
- [15] VAPNIK V N, KOTZ S. Estimation of dependences based on empirical data[M]. New York: Springer, 2006.
- [16] KIM W, KIM K S, LEE J E, et al. Development of novel breast cancer recurrence prediction model using support vector machine[J]. *Journal of Breast Cancer*, 2012, 15(2): 230-238.
- [17] DEPEURSINGE A, KURTZ C, BEAULIEU C, et al. Predicting visual semantic descriptive terms from radiological image data: preliminary results with liver lesions in CT[J]. *IEEE Transactions on Medical Imaging*, 2014, 33(8): 1669-1676.
- [18] 王颖霖, 郭伟, 师越, 等. 基于 Grey-Sigmoid 核函数支持向量机高光谱遥感图像分类方法研究[J]. *河南农业大学学报*, 2017, 51(3): 402-407.
- WANG H L, GUO W, SHI Y, et al. Study on hyperspectral remote sensing image classification method based on the grey-sigmoid kernel function support vector machine[J]. *Journal of Henan Agricultural University*, 2017, 51(3): 402-407.
- [19] 徐存东, 张锐, 王荣荣, 等. 基于改进支持向量机的盐碱地信息精确提取方法研究[J]. *灌溉排水学报*, 2018, 37(9): 62-68.
- XU D D, ZHANG Y, WANG R R, et al. An Improved Support Vector Machine Method for Estimating Saline-alkali Soil from Remote Sensing Imagery[J]. *Journal of Irrigation and Drainage*, 2018, 37(9): 62-68.
- [20] FARHIDZADEH H, GOLDFOG D B, HALL L O, et al. Texture feature analysis to predict metastatic and necrotic soft tissue sarcomas [C]//2015 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE press, 2015: 2798-2802.

Prediction of soft tissue sarcoma metastasis by PET-CT image texture features

Shen Junli, Yu Kun

(School of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China)

Abstract: This paper proposes an auxiliary diagnostic method for soft tissue sarcoma metastasis prediction. This method extracts 105 features which include 24 features of the Gray Level Co-occurrence Matrix (GLCM) and other 81 grayscale features by analyzing the texture features of FDG-PET and CT diagnostic images. Machine learning algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Random Forest (RF) are used to build prediction models, and their parameters are optimized by grid search method. Finally, the models are evaluated by the leave-one-out cross-validation method. By evaluating the performance of each model, support vector machine can be selected as the final prediction model, and the average accuracy of 80% is obtained. In addition, the sensitivity and specificity of this model reached 81% and 79% respectively, indicating that the predicted results of this model have certain reliability, which can be used to aid diagnosis of STS and improve patient outcomes through better adaptive treatment.

Keywords: soft tissue sarcoma; texture feature; machine learning; metastatic prediction

[责任编辑 陈留院 赵晓华]