DOI: 10.16781/j.CN31-2187/R.20230560



## 基于影像组学分析的椎弓根螺钉钉道超声智能评估

唐天灵<sup>1</sup>, 马烨波<sup>1</sup>, 杨 桓<sup>2</sup>, 叶长青<sup>1</sup>, 孔佑进<sup>1</sup>, 杨卓畅<sup>1</sup>, 周 昌<sup>1</sup>, 邵 杰<sup>2</sup>, 孟炳堃<sup>2</sup>, 王卓然<sup>1</sup>, 陈建刚<sup>1,3\*</sup>, 陈自强<sup>2\*</sup>

1. 华东师范大学上海市多维度信息处理重点实验室, 上海 200241

2. 海军军医大学(第二军医大学)第一附属医院脊柱外科,上海 200433

3. 上海中医药大学康复医学院中医智能康复教育部工程研究中心, 上海 201203

[摘要] **目** 69 基于影像组学分析提出一种针对椎弓根螺钉钉道超声图像的分类方法,以对钉道完整性进行评估。方法 利用 4 例新鲜尸体的胸腰椎标本预建立 50 个钉道并获取钉道超声图像,选取 2 000 张图像(钉道完整与破损的样本各 1 000 个),采用五折交叉验证的方法将数据集按照 4 : 1 的比例进行随机划分,形成训练集和测试集(分别包含 1 600 个和 400 个样本)。首先利用大津阈值法找到感兴趣区的最佳半径,然后用 pyradiomics 提取组学特征,再采用主成分分析算法和最小绝对收缩和选择算子算法分别进行降维和特征筛选,最后分别使用支持向量机(SVM)、logistic 回归、随机森林 3 种机器学习模型和视觉几何组网络(VGG)、残差网络(ResNet)、转换器模型(Transformer)3 种深度学习模型对超声图像进行分类。采用准确度对各模型的分类性能进行评估。结果 在感兴趣区半径为 230 像素时,SVM 模型的分类准确度最高,为 96.25%;而 VGG 模型的准确度只有 51.29%, logistic 回归、随机森林、ResNet、Transformer 模型的准确度分别为 85.50%、80.75%、80.17%、75.18%。结论 在对椎弓根螺钉钉道超声图像的分类方面,机器学习模型整体上相较于深度学习模型表现更好,其中 SVM 模型的分类性能最佳,可用于辅助医师诊断。

[关键词] 椎弓根螺钉置入; 超声检查; 影像组学; 支持向量机; 机器学习; 人工智能

[引用本文] 唐天灵,马烨波,杨桓,等.基于影像组学分析的椎弓根螺钉钉道超声智能评估[J].海军军医大学学报,2024,45(11):1362-1370.DOI: 10.16781/j.CN31-2187/R.20230560.

#### Intelligent assessment of pedicle screw canals with ultrasound based on radiomics analysis

TANG Tianling<sup>1</sup>, MA Yebo<sup>1</sup>, YANG Huan<sup>2</sup>, YE Changqing<sup>1</sup>, KONG Youjin<sup>1</sup>, YANG Zhuochang<sup>1</sup>, ZHOU Chang<sup>1</sup>, SHAO Jie<sup>2</sup>, MENG Bingkun<sup>2</sup>, WANG Zhuoran<sup>1</sup>, CHEN Jiangang<sup>1,3\*</sup>, CHEN Ziqiang<sup>2\*</sup>

 Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China
 Department of Spinal Surgery, The First Affiliated Hospital of Naval Medical University (Second Military Medical University), Shanghai 200433, China

3. Engineering Research Center of Traditional Chinese Medicine Intelligent Rehabilitation, Ministry of Education, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

[Abstract] Objective To propose a classification method for ultrasound images of pedicle screw canals based on radiomics analysis, and to evaluate the integrity of the screw canal. Methods With thoracolumbar spine specimens from 4 fresh cadavers, 50 pedicle screw canals were pre-established and ultrasound images of the canals were acquired. A total of 2 000 images (1 000 intact and 1 000 damaged canal samples) were selected. The dataset was randomly divided in a 4 : 1 ratio using 5-fold cross-validation to form training and testing sets (consisting of 1 600 and 400 samples, respectively). Firstly, the optimal radius of the region of interest was identified using the Otsu's thresholding method, followed by feature extraction

#### [收稿日期] 2023-10-10 [接受日期] 2024-08-26

[基金项目] 国家自然科学基金(82151318),上海市自然科学基金(21ZR1478600),上海市科技计划项目(21Y11902500),2023 年地方科学发展基金(XZ202301YD0032C),吉林省科技发展计划项目(20230204094YY),2022 年度教育部 "春晖计划"合作科研项目,海军军医大学(第二军 医大学)"远航"计划. Supported by National Natural Science Foundation of China (82151318), Natural Science Foundation of Shanghai (21ZR1478600), Shanghai Science and Technology Project (21Y11902500), 2023 Local Science Development Fund (XZ202301YD0032C), Science and Technology Development Project of Jilin Province (20230204094YY), 2022 "Chunhui" Cooperation Project of The Ministry of Education, and "Long Voyage" Project of Naval Medical University (Second Military Medical University).

[作者简介] 唐天灵,硕士生.E-mail: 51255904077@stu.ecnu.edu.cn

\*通信作者(Corresponding authors). Tel: 021-54345193, E-mail: gchen@cee.ecnu.edu.cn; Tel: 021-31161697, E-mail: zigiang\_chensuper81@vip.163.com

using pyradiomics. Principal component analysis and the least absolute shrinkage and selection operator algorithm were employed for dimensionality reduction and feature selection, respectively. Subsequently, 3 machine learning models (support vector machine [SVM], logistic regression, and random forest) and 3 deep learning models (visual geometry group [VGG], ResNet, and Transformer) were used to classify the ultrasound images. The performance of each model was evaluated using accuracy. **Results** With a region of interest radius of 230 pixels, the SVM model achieved the highest classification accuracy of 96.25%. The accuracy of the VGG model was only 51.29%, while the accuracies of the logistic regression, random forest, ResNet, and Transformer models were 85.50%, 80.75%, 80.17%, and 75.18%, respectively. **Conclusion** For ultrasound images of pedicle screw canals, the machine learning model performs better than the deep learning model as a whole, and the SVM model has the best classification performance, which can be used to assist physicians in diagnosis.

[Key words] pedicle screw implantation; ultrasonography; radiomics; support vector machine; machine learning; artifical intelligence

[Citation] TANG T, MA Y, YANG H, et al. Intelligent assessment of pedicle screw canals with ultrasound based on radiomics analysis[J]. Acad J Naval Med Univ, 2024, 45(11): 1362-1370. DOI: 10.16781/j.CN31-2187/R.20230560.

椎弓根螺钉技术可有效应用于脊柱重建<sup>[1]</sup>,是 目前治疗脊柱畸形、脊柱肿瘤、脊柱骨折和脊柱退 行性疾病的最常见手术术式<sup>[23]</sup>。然而,人体脊柱是 一个复杂的三维结构,其周围存在许多重要组织, 如脊髓、神经根、大血管和胸腹部脏器等,且在颈 椎、胸椎等脊柱高位节段的椎弓根较细,准确置入 手术的操作难度较大。手术过程中螺钉误置可能引 起出血、脏器损伤,甚至可导致瘫痪和死亡<sup>[4]</sup>。导 航系统近年来发展迅速,可用于辅助椎弓根螺钉置 入,帮助医师在手术过程中更准确地定位和置入椎 弓根螺钉,从而减少手术时间和风险,并提高手术 的成功率。

目前,可用的椎弓根螺钉辅助置入导航系统 主要是基于 CT 影像配准的导航系统<sup>[5-8]</sup>,然而 CT 成像存在潜在有害的电离辐射,并且会增加手术 时间。与CT、MRI相比,超声检查经济便捷,还 有着不劣于 CT 的诊断性能<sup>[9]</sup>。因此,一些基于超 声的导航系统应运而生<sup>[10]</sup>,未来可能成为替代术 中CT和X线透视的低成本和无辐射方案。本课题 组既往研究探讨了一种基于支持向量机(support vector machine, SVM)构建的椎弓根螺钉钉道完 整性超声图像鉴别与验证方法<sup>[11]</sup>。该方法不仅具 有超声操作简单、实时性高、无电离辐射的优点, 也利用计算机自动辅助判断钉道完整性的功能方便 了超声技术的临床应用,无须依赖专业医师对图像 进行判断和解释。但是,此方法仅基于人工定义的 少量特征,缺乏对图像多种特征的提取与分析。本 研究在前期工作的基础上,基于更大的数据集,采 用 pyradiomics 提取大量影像组学特征,使用 SVM

进行分类,并与多种机器学习和深度学习分类器的 分类结果进行比较,现将结果报告如下。

#### 1 材料和方法

1.1 材料与仪器 4例新鲜尸体的胸腰椎标本 由海军军医大学(第二军医大学)基础医学院解 剖学教研室提供,男、女各2例,年龄为59~85 (73.14±9.87)岁。本实验获得海军军医大学(第 二军医大学)第一附属医院伦理委员会审批。预建 立钉道共50个,使用20 MHz超声换能器(型号 为Olympus UM-S20-17S,探头频率为20 MHz、直 径为1.7 mm;日本Olympus公司)对预建立的钉道 底部进行超声检查并采集图像,使用CT机(型号为 Aquilion ONE,日本Canon 医疗系统有限公司)对预 建立的钉道进行完整性评估。使用基于Windows 11 系统的16核3.20 GHz CPU、16 GB 内存的计算机 并搭配 GPU GTX3060 进行模型训练和预测。

1.2 研究方法 整体流程图见图 1。由 3 名至少有 5 年脊柱外科工作经验的医师(包括 2 名主治医师 和 1 名副主任医师)对钉道完整性进行判读,并采 用实际探测、CT 图像和钉道超声图像相结合的方 式对超声图像进行分类。分类任务首先由 2 名医师 各自独立完成,若 2 人意见存在分歧,则由第 3 名 医师进行裁决。在分类过程中,任何存在重叠或不 清晰的超声图像均被排除。将医师判定钉道完整无 破损的图像定义为阴性样本,将医师判定钉道破损 的图像定义为阳性样本。最终用于研究的数据集包 括阴性样本和阳性样本各 1 000 个,共 2 000 个样 本。本研究按4 : 1 的比例将数据集随机划分训练



集和测试集(分别包含1600个和400个样本),

并进行五折交叉验证。



A: Preprocessed image; B: Foreground and background segmented image by OTSU; C: Mask. OTSU: Otsu's thresholding method; PCA: Principal component analysis; LASSO: Least absolute shrinkage and selection operator; SVM: Support vector machine; VGG: Visual geometry group; ROC: Receiver operating characteristic.

采用 pyradiomics 对超声图像提取组学特征, 用SVM模型对图像进行钉道是否破损的二分类判 断。阴性样本和阳性样本如图2所示,阴性样本的 钉道内部有1个较亮的圆环,通常被称为钉道环或 钉道声影, 它是由钉道周围的软组织及钉道内的物 质反射和散射超声波而形成的。阳性样本的钉道 内部通常不会形成较亮的圆环,这是因为钉道破损 后,钉道壁和内容物发生了变化,导致超声波传播 的路径发生改变,从而影响了超声波的反射和散 射,使超声波在钉道内部形成弱反射或散射,形成 1个较暗的区域或较亮的斑点,导致钉道环消失或 变得模糊。同时, 阴性样本由于钉道没有破损, 声 波不会通过破损的钉道壁在钉道外形成声波反射, 即钉道外部不会产生白色前景。而在阳性样本中, 声波会通过破损的钉道壁在钉道外形成声波反射, 即钉道外部会产生白色前景。

基于 pyradiomics 的纹理特征分析有可能将阴 性样本与阳性样本区分开。首先使用大津阈值法 (Otsu's thresholding method, OTSU)得到用于区 分前景和背景的阈值 *T*, 然后做椎弓根螺钉钉道的 同心圆(直径为椎弓根螺钉钉道的直径),该同心圆的判别半径是根据钉道外部的特征并以此设计的损失函数得到的。考虑到基层医疗机构较低配置的需求同时加快运算分析速度,本研究仅提取同心圆内部图像的特征并进行降维筛选,最后使用 SVM进行分类。





Negative samples have a bright circular ring inside the nail. Positive samples do not form a bright circular ring, but rather a darker area or brighter spots.

1.2.1 图像预处理 (1)图像裁剪: 原始超声图像包含了许多非必要的信息, 如患者信息和超声仪设备名等。为了避免这些额外信息对结果产生影响, 同

时尽可能减轻处理器的算力负担,需要对原始超声图像进行裁剪,将尺寸为1920像素×1080像素的原

始图像(图 3A) 裁剪为尺寸为 894 像素 ×764 像素 的图像(图 3B)。



## 图 3 图像预处理 Fig 3 Image preprocessing

A: Original image; B: Cropped image; C: Enhanced image; D: Center position of undamaged spine on ultrasound image. The center coordinates are (505, 383).

(2)图像增强:为了改善图像质量,提高图像中前景和背景的区分度,改善图像判读效果,使用灰度变换函数将图像I进行灰度映射变换。该变换能够改变图像的整体亮度和对比度并增强图像中暗部细节的可见度,从而改善图像的整体可见性和质量。将增强后的图像记为J,即

(3)确定圆心坐标:由于该数据集来自同一台 超声设备并且设置保持不变,因此数据集中所有超声 影像的圆心坐标相同(图 3D)。设图像左上角为原 点,坐标为(0,0),圆心坐标均为(505,383)。 1.2.2 特征提取 为了在保证分类性能的前提下使 该方法适用于轻量化设备,同时加快分析速度,仅 对钉道内部图像进行特征提取及分析。因此,采用 钉道的同心圆作为所有样本的感兴趣区(region of interest, ROI)并进行特征提取。

(1)阈值 T 的选取:利用 OTSU 计算 ROI 的 具体半径。OTSU 是一种确定图像二值化分割阈值 的算法,由日本学者大津于 1979 年提出<sup>[12]</sup>。根据图 像的灰度特性,通过计算前景和后景最大类间方差 求其阈值。首先,通过图像的灰度直方图,计算每个 像素在整幅图像的概率分布。再根据灰度直方图中 的 2 个峰值,假设阈值 T 并将所有像素分为 2 类,分 别是小于 T 的 C<sub>1</sub> 和大于 T 的 C<sub>2</sub>。图像全局均值 M<sub>G</sub> 可由下列公式求得

$$P_1 = \sum_{i=0}^{k} p_i \tag{2}$$

$$M_{1} = \frac{\sum_{i=0}^{k} ip_{i}}{P_{1}} = \frac{\sum_{i=0}^{k} ip_{i}}{\sum_{i=0}^{k} p_{i}}$$
(3)

$$M_{2} = \frac{\sum_{i=k+1}^{L-1} ip_{i}}{P_{2}} = \frac{\sum_{i=k+1}^{L-1} ip_{i}}{\sum_{i=k+1}^{L-1} p_{i}}$$
(4)

$$M_G = P_1 M_1 + P_2 M_2 \tag{5}$$

其中 $P_1$ 、 $P_2$ 分别是像素被分为 $C_1$ 、 $C_2$ 的概率,  $M_1$ 、 $M_2$ 分别是像素集 $C_1$ 、 $C_2$ 各自的均值, $p_i$ 是灰 度值为*i*时的概率,*k*是所求阈值的灰度值,*L*是图 像的最大灰度值。且

$$P_1 + P_2 = 1$$
 (6)

然后求类间方差 o<sup>2</sup>, 表达式为

 $\sigma^{2} = P_{1} (M_{1} - M_{G})^{2} + P_{2} (M_{2} - M_{G})^{2}$ (7) 上式可以化简为

$$\sigma^2 = P_1 P_2 (M_1 - M_2)^2 \tag{8}$$

最后,只要遍历灰度值 0~255,使 σ<sup>2</sup> 最大的灰度 值 k 就是大津阈值 T,即用于判断某像素属于前景 还是后景的阈值 T。

(2)半径 R 的选取:对于数据集中的所有超 声图像,以钉道中心(505,383)为圆心作半径为 R 的同心圆。对于所有阴性样本,统计半径为 R 时 圆外存在像素点灰度值大于阈值 T 的图像(即同心 圆外存在前景的图像)数量,记为 N<sub>1</sub>。对于所有阳 性样本,统计圆外存在像素点灰度值大于阈值 T 的 图像数量,记为 N<sub>2</sub>。令 N 为 N<sub>1</sub>、N<sub>2</sub> 差值的绝对值

N=|N<sub>1</sub>-N<sub>2</sub>| (9) 在不同半径下,当N取到最大值时表明此时的半径 对该数据集钉道完整性判断的准确度达到最佳。 (3)组学特征提取:采用 pyradiomics 提取超 声图像 ROI 的组学特征。pyradiomics 是一个开源 的 Python 库,能够利用灰度共生矩阵、灰度依赖矩 阵、灰度大小共生矩阵等特征提取方法从医学图像 中提取大量的影像组学特征。这些特征包括一阶特 征、形状特征、灰度共生矩阵特征、灰度大小区域 矩阵特征、灰度游程矩阵特征、相邻灰度色调差异 矩阵特征和灰度依赖矩阵特征等。这些特征可用于 肿瘤诊断、治疗选择、预后评估等。

1.2.3 特征降维和筛选 在机器学习中,特征降 维和筛选是非常重要的技术,它们的主要目的是 减少特征数量,提高模型的性能和效率。固然 pyradiomics 提供了大量的有用特征, 然而其中很多 特征可能冗余或不相关,这些特征会导致模型过拟 合、训练时间长、预测效率低等。特征降维和筛选 可以使模型更加关注数据中最重要的特征,从而提 高模型的准确度和泛化能力。特征降维和筛选还 可以将高维数据降低到二维或三维,从而更容易对 数据进行可视化和分析。通过可视化可以发现数 据中的特殊规律和关系,为后续的建模和预测提供 更好的启示。本研究对特征进行降维和筛选处理, 既保证了模型性能,也符合轻量化的要求。本研究 选取了主成分分析 (principal component analysis, PCA)算法进行特征降维,最小绝对收缩和选择算  $\neq$  (least absolute shrinkage and selection operator, LASSO)算法进行特征筛选。

PCA 是一种常见的数据降维技术,通过线性变 化将高维数据映射到低维空间,找到能够最大程度 保留原始数据信息的新特征表示。PCA 通过线性 变换将原始数据投影到一个新的坐标系中,使数据 在新坐标系中的方差最大化。这些新的坐标轴被称 为主成分,每个主成分代表原始数据中的一部分信 息。通俗来讲,就是将特征按一定的比例压缩。

LASSO 是一种常用的线性回归正则化方法, 它通过对模型参数施加 L1 正则化来实现特征筛选 和降维。相较于传统的线性回归方法, LASSO 能 够更好地处理高维数据,避免数据过拟合。LASSO 的代价函数 *J*(*θ*) 为

 $J(\theta) = \frac{1}{2m} \left[ \sum_{i=1}^{m} (h_{\theta}(x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^{n} |\theta_j| \right] (10)$ 1.2.4 分类模型 本研究旨在运用机器学习技术 进行椎弓根螺钉钉道超声图像二分类,判断椎弓根 螺钉钉道是否破损,从而辅助医师进行诊断。为 了凸显影像组学和机器学习相结合的优势,在经 典机器学习分类方法中选择最佳的分类器,本研 究选取3种模型进行比较,分别为SVM、随机森 林、logistic 回归模型。此外,为了阐明机器学习 小样本数据集中的优越性,本研究对比了3种经典 的深度学习模型,分别是视觉几何组网络(visual geometry group, VGG)、残差网络(ResNet)、 转换器(Transformer)模型。

深度学习分析的具体流程和参数如下:(1)在 对图像输入模型进行训练及预测前,统一调整图像 大小为 224 像素×224 像素。(2)选用交叉熵损 失函数计算模型预测值与真实值的差值。(3)考 虑到计算机内存、GPU、数据集的样本大小和输入 图像大小,设置批大小(BacthSize)为16。所有 模型均在训练30轮后停止训练,即在训练完成前 所有图像都被迭代了30次。

SVM 是一种常用的监督学习算法,通常用于 分类和回归任务。SVM 的目标是找到一个最优的 超平面(或多个平面),将不同类别的数据分开, 使不同类别的数据能够被最大程度地分离,同时最 小化分类错误的概率。SVM 算法的核心思想是通 过寻找最大化边缘(即最大化不同类别之间的最小 距离)的超平面来进行分类,划分超平面可定义为 一个线性方程,即

w<sup>T</sup>X+b=0 (11)
式中w={w<sub>1</sub>;w<sub>2</sub>;…;w<sub>d</sub>},是一个法向量,决定了超平面的方向,T表示法向量w的转置,X为训练样本,b为位移项(决定了超平面与原点之间的距离)。
1.3 统计学处理采用准确度、特异度、灵敏度(召回率)、F1值对分类结果进行评估,这些评估指标基于每次交叉验证实验所得结果的平均值计算得到,分别为

准确度=
$$\frac{TP+TN}{TP+TN+FN+FP}$$
×100% (12)

特异度=
$$\frac{1N}{TN+FP}$$
×100% (13)

灵敏度(召回率) = 
$$\frac{\mathrm{IP}}{\mathrm{TP} + \mathrm{FN}} \times 100\%$$
 (14)

精确率=
$$\frac{1P}{TP+FP}$$
×100% (15)

F1 值=2×
$$\frac{桐硼率 × 召回率}{精确率+召回率}$$
×100% (16)

其中TP为真阳性例数,TN为真阴性例数,FP为 假阳性例数,FN为假阴性例数。 同时采用ROC曲线和精确率-召回率 (precision-recall, PR)曲线对6种模型的分类性 能进行分析和评估,利用AUC值评价模型的分类 性能。

#### 2 结 果

2.1 最佳ROI半径与关键特征的确定 涌过 OTSU求得区分超声图像前后景的阈值T为48。 如图 4A 所示, 在此阈值条件下, 计算得到 N 的最 大值为 961, 此时 ROI 的最佳半径 R 为 230 像素。 采用 pyradiomics 提取到组学特征共1316个。利 用PCA算法将提取到的特征降维到150维,通过 LASSO算法将特征筛选到8个(图4B)。这8个 特征所对应的贡献率最高的原始特征如表1所 示, wavelet-HHL gldm SmallDependenceEmphasis 描述了图像纹理的小依赖程度,即图像中微小 结构的分布和依赖关系;当钉道存在微小破损 时,图像中的微小结构分布可能会发生变化。 wavelet-LLH ngtdm Contrast 和 wavelet-HHH gldm GrayLevelVariance分别描述了图像纹理的 对比度和灰度级方差: 当存在破损情况时, 破损 区域与周围正常组织之间的灰度差异可能更加明 显,从而影响到这2个特征的值。wavelet-HLL glrlm LongRunHighGrayLevelEmphasis 和 wavelet-HHH glrlm LowGrayLevelRunEmphasis 分别描述 了图像高、低灰度级的强调度,即图像中存在的 灰度级较高和较低部分的纹理特征的程度:在 椎弓根钉道超声图像中,这可能与图像中的前景 和背景的纹理特征有关。wavelet-HHL glszm SizeZoneNonUniformity 和 wavelet-HHH gldm GrayLevelNonUniformity 分别描述了图像中不同尺 寸区域和灰度级的非均匀性,不同尺寸区域的非均 匀性可能反映了椎弓根钉道中不同区域的分布和形 态特征, 灰度级的非均匀性可能反映了椎弓根钉道 中不同组织之间的灰度差异和纹理特征的变化。 squareroot ngtdm Coarseness 描述了图像的粗糙 度,破损椎弓根钉道超声图像纹理的不规则状态更 为显著,粗糙度值可能更高。由此可见,这8个特 征涵盖了图像纹理的不同方面,包括纹理依赖度、 对比度、灰度级别方差、长运行高灰度级别强调 度、大小区域非均匀性、低灰度级别运行强调度、 粗糙度和灰度级别非均匀性。







A: Selection of ROI radius *R*. When the *N* reaches its maximum (961), the optimal radius of ROI was 230 pixels; B: LASSO feature selection. The vertical axis represents the feature No. after dimensionality reduction using PCA, while the horizontal axis represents the LASSO feature coefficient. ROI: Region of interest; LASSO: Least absolute shrinkage and selection operator; PCA: Principal component analysis.

本研究利用OTSU确定了ROI半径。为了验证 230 像素是否为最佳半径,选取了150、200、210、 220、230、240、250 像素作为ROI半径,采用影像 组学和SVM组合模型分别进行实验。不同半径下 模型的分类性能见图5,由图可见分类性能在ROI 半径为150~230 像素时呈上升趋势,在230 像 素达到顶峰,而在230~250 像素时又呈现下降趋势。结果表明随着半径的扩大,ROI包含的有效信息增多,而在230 像素之后,半径越大也代表着包含的冗余信息越多,干扰了模型的判断。因此, 230 像素是ROI的最佳半径,而且用OTSU计算最 佳半径是可行的。

## 表 1 LASSO 算法筛选的 8 个特征所对应的贡献率 最高的原始特征

# Tab 1Original feature with the highest contributionrate corresponding to features selected by LASSO

FN	Original feature with the highest contribution rate
2	wavelet-HHL_gldm_SmallDependenceEmphasis
6	wavelet-LLH_ngtdm_Contrast
3	wavelet-HHH_gldm_GrayLevelVariance
23	wavelet-HLL_glrlm_LongRunHighGrayLevelEmphasis
37	wavelet-HHL_glszm_SizeZoneNonUniformity
5	wavelet-HHH_glrlm_LowGrayLevelRunEmphasis
26	squareroot_ngtdm_Coarseness
10	wavelet-HHH gldm GrayLevelNonUniformity

LASSO: Least absolute shrinkage and selection operator; FN: Feature No.

2.2 模型分类性能对比 在 230 像素的半径下, 各分类器对测试集的分类性能如表 2 所示。在测试 集中, SVM模型的分类准确度为 96.25%, logistic 回归模型的准确度为 85.50%,随机森林模型的准 确度为 80.75%, VGG 模型的准确度为 51.29%, ResNet 模型的准确度为 80.17%, Transformer 模型 的准确度为 75.18%。



ROI radii

As the radius increases, the accuracy of the SVM model gradually increases. It reaches its highest point at a radius of 230 pixels, with an accuracy of 96.25%. However, as the radius further increases, the accuracy starts to decline. ROI: Region of interest; SVM: Support vector machine.

## 表 2 各模型测试集分类性能 Tab 2 Classification performance of each model on test set

				(%), $n = 400, \bar{x} \pm s$
Model	Accuracy	Specificity	Sensitivity	F1 score
Machine learning				
SVM	$96.25 \pm 0.85$	$97.00 \pm 1.22$	$95.50 \pm 1.14$	$96.22 \pm 0.86$
Logistic regression	$85.50 \pm 1.13$	$89.70 \pm 1.03$	$81.30 \pm 1.36$	$84.86 \pm 1.21$
Random forest	$80.75 \pm 0.35$	$90.10 \pm 1.46$	$71.40 \pm 1.66$	$78.76 \pm 0.60$
Deep learning				
VGG	$51.29 \pm 0.27$	$98.79 \pm 1.20$	$3.79 \pm 1.00$	$7.19 \pm 1.76$
ResNet	$80.17 \pm 0.43$	$65.78 \pm 1.30$	$94.57 \pm 1.04$	$82.67 \pm 0.36$
Transformer	$75.18 \pm 1.11$	$52.94 \pm 1.95$	$96.48 \pm 1.23$	$79.89 \pm 0.84$

SVM: Support vector machine; VGG: Visual geometry group.

本研究使用 AUC 值作为评估指标,对分类器 性能进行全面的定量评估。ROC 曲线(图 6A) 分析结果显示, SVM 模型分类椎弓根螺钉钉道 是否破损的 AUC 值为 0.97, logistic 回归模型的 AUC 值为 0.80,随机森林模型的 AUC 值为 0.95, ResNet 模型的 AUC 值为 0.92, VGG 模型的 AUC 值为 0.63, Transformer 模型的 AUC 值为 0.77。各 模型的 PR 曲线如图 6B 所示,使用 PR 曲线的 AUC 值作为度量指标,SVM 模型的 AUC 值为 0.97, logistic 回归模型的 AUC 值为 0.87,随机森林模型 的 AUC 值为 0.96, ResNet 模型的 AUC 值为 0.93, VGG 模型的 AUC 值为 0.60, Transformer 模型的 AUC 值为 0.80。由此可见, SVM 和随机森林模型 具有较高的分类性能, 而 VGG 模型的分类性能相 对其他模型较差; 综合考虑 6 种模型的分类结果, 机器学习模型整体表现优于深度学习模型。

2.3 模型复杂度分析 从模型的算法复杂度来 看,本研究训练集的样本量为1600例,机器学习 模型的输入特征维度均为8维。SVM模型的算法 复杂度主要与支持向量的数量相关,本实验所用模 型的支持向量数量为537。logistic回归是一种线 性模型,其算法复杂度主要与特征维度相关。随机 森林是一种集成学习方法,它由多个决策树组成, 算法复杂度主要与决策树的数量相关,本研究中决 策树的数量为30。本研究所用深度学习模型的参数量和计算量见表3,可以看出深度学习的算法复杂度远高于机器学习模型。因此,相对深度学习模

型,机器学习模型的算法复杂度整体较低,而在分 析钉道超声图像时仍取得了较好的性能,更适用于 本研究的数据集。





ROC: Receiver operating characteristic; PR: Precision-recall; SVM: Support vector machine; LR: Logistic regression; RF: Random forest; VGG: Visual geometry group; AUC: Area under curve.

### 表 3 各深度学习模型参数量和计算量

Tab 3 Parameters and FLOPs of each deep

learning model						
Model	Parameter ( $\times 10^{6}$ )	FLOPs ( $\times 10^9$ )				
ResNet	25.6	4.13				
VGG	138	15.5				
Transformer	8.82	4.37				

FLOPs: Floating point operations; VGG: Visual geometry group.

#### 3 讨 论

本研究基于影像组学提出了一种针对椎弓根螺 钉钉道超声图像的分类方法,对钉道完整性进行评 估。该方法首先利用OTSU找到ROI的最佳半径, 再用pyradiomics提取组学特征,最后采用PCA算 法和LASSO算法进行降维和特征筛选。本研究将 钉道图像按4:1划分为训练集和测试集,分别使 用SVM、logistic回归、随机森林3种机器学习模 型和VGG、ResNet、Transformer3种深度学习模 型对超声图像进行分类。实验结果表明,SVM模 型的准确度最高。从全局来看,机器学习模型的性 能整体优于深度学习模型。其原因在于机器学习比 深度学习更适用于小样本训练。在本实验中SVM 的表现优于其他机器学习模型,其原因主要有: (1) SVM 可以在输入空间中构建一个超平面来分割不同的类别,有效地处理高维数据; (2) SVM 通过最大化间隔来确定决策边界,使得其在训练过程中对噪声和异常点具有很好的鲁棒性。

在实际情况中,仅凭借一张图像无法可靠地判 断钉道是否完整。判断钉道情况需要对该患者钉道 进行不同深度的超声扫描,并结合专业医师的判断 和分析才能得出可靠的结论。而本研究方法可以作 为一种辅助手段,通过基于影像组学的图像分类方 法,能够有效地提高钉道图像完整性评估的准确性 和效率。与传统的手动分析方法相比,本方法具有 更高的自动化程度和更快的处理速度,可以使医师 更加专注于对图像结果的分析和判断。此外,本方 法还可以通过大量的数据训练和特征降维,进一步 提高图像分类的准确性和鲁棒性,为医师进行钉道 超声图像判断和诊断提供更加可靠和稳定的支持。

在前期研究中,本课题组提出了一种基于超 声图像智能分析的椎弓根螺钉钉道完整性评估方 法<sup>[11]</sup>,该方法首先将纹理特征作为第一类特征 对超声图像特征作初始分类,筛选出椎弓根螺钉 钉道破损严重的超声图像,再选取熵、方差、对 比度、能量、平均绝对偏差5个参数作为第二类 特征,对超声图像进行二次分类。该实验结果表 明,在采集的800个样本中,二次分类的总体准 确度为94.25%。本研究方法较前期研究有所改 进,具体为:(1)不再手动提取特征,而是采用 pyradiomics 提取影像组学特征,大大提高了获取特 征的数量和质量; (2)扩大了数据量,从800个 样本提升到2000个样本; (3)提高了准确度, 准确度从之前的94.25%提高到96.25%; (4)与 其他分类器进行了比较,本研究比较了3个机器学 习分类器和3个深度学习分类器。

本研究存在以下局限性: (1)数据集共入选 2 000 个样本, 样本量基本满足机器学习的要求, 然而,当将这些数据应用于深度学习时模型的分类 效果并不理想, 表明样本量仍需进一步扩大。未来 将尝试增加更多的数据样本,进一步提高深度学习 模型的分类效果。(2)虽然使用了现有的分类模 型和降维算法进行图像分类,但仍存在不足。例如, 在某些情况下,图像数据可能会受到噪声、伪影等 因素的影响,从而导致分类结果的不准确性。为了 克服这些不足,未来的研究可以尝试改进分类模型 和降维算法,提高自适应能力和分类精度,以便更好 地适应不同类型的图像数据和特征,进而提高基于 影像组学的图像分类方法的准确性和可靠性。

综上所述,本研究提出的基于影像组学的 SVM模型有望对椎弓根螺钉钉道是否破损进行准 确分类,其可辅助医师对钉道超声图像进行判断和 解释, 尤其可以辅助经验欠缺的医师。

## 「参考文献]

- [1] GAINES R W Jr. The use of pedicle-screw internal fixation for the operative treatment of spinal disorders[J]. J Bone Joint Surg Am, 2000, 82(10): 1458-1476. DOI: 10.2106/00004623-200010000-00013.
- [2] PERNA F, BORGHI R, PILLA F, et al. Pedicle screw insertion techniques: an update and review of the literature [J]. Musculoskelet Surg, 2016, 100(3): 165-169. DOI: 10.1007/s12306-016-0438-8.
- 范少勇,胡梁深,华鸿,等.计算机辅助导航技术在胸 [3] 腰椎椎弓根螺钉内固定术的临床应用[J]. 江西中医 药,2023,54(6):44-45,48.

- [4] DEDE O, WARD W T, BOSCH P, et al. Using the freehand pedicle screw placement technique in adolescent idiopathic scoliosis surgery: what is the incidence of neurological symptoms secondary to misplaced screws?[J]. Spine (Phila Pa 1976), 2014, 39(4): 286-290. DOI: 10.1097/ BRS.00000000000127.
- [5] LARSON A N, POLLY D W Jr, GUIDERA K J, et al. The accuracy of navigation and 3D image-guided placement for the placement of pedicle screws in congenital spine deformity[J]. J Pediatr Orthop, 2012, 32(6): e23-e29. DOI: 10.1097/BPO.0b013e318263a39e.
- [6] MASON A, PAULSEN R, BABUSKA J M, et al. The accuracy of pedicle screw placement using intraoperative image guidance systems [J]. J Neurosurg Spine, 2014, 20(2): 196-203. DOI: 10.3171/2013.11.SPINE13413.
- [7] 胡定祥,郑瑞清,李长辉,等.3D打印导航模块在胸 腰椎椎弓根螺钉精准置入中的应用价值[J].解剖 学报,2023,54(3):342-347. DOI: 10.16098/j.issn. 0529-1356.2023.03.013.
- [8] MALHAM G M, MUNDAY N R. Comparison of novel machine vision spinal image guidance system with existing 3D fluoroscopy-based navigation system: a randomized prospective study [J]. Spine J, 2022, 22(4): 561-569. DOI: 10.1016/j.spinee.2021.10.002.
- [9] 毛玉梅,张雅丽.肌骨超声检查与CT、MRI用于创 伤性浅表软组织损伤的诊断价值比较[J].影像研 究与医学应用,2023,7(8):166-169. DOI: 10.3969/ j.issn.2096-3807.2023.08.054.
- [10] GUEZIRI H E, GEORGIOPOULOS M, SANTAGUIDA C, et al. Ultrasound-based navigated pedicle screw insertion without intraoperative radiation: feasibility study on porcine cadavers[J]. Spine J, 2022, 22(8): 1408-1417. DOI: 10.1016/j.spinee.2022.04.014.
- [11] 马烨波,邵杰,杨桓,等.基于机器学习的椎弓根螺 钉钉道超声图像评估方法[J].海军军医大学学报, 2022,43(9):993-999. DOI: 10.16781/j.CN31-2187/ R.20220321.

MA Y B, SHAO J, YANG H, et al. Evaluation of ultrasonic images of pedicle screw channel based on machine learning[J]. Acad J Naval Med Univ, 2022, 43(9): 993-999. DOI: 10.16781/j.CN31-2187/R.20220321.

[12] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Trans Syst Man Cybern, 1979, 9(1): 62-66. DOI: 10.1109/TSMC.1979.4310076.

[本文编辑] 杨亚红