

机器学习在脊柱疾病智能诊治中的应用综述

刘彤彤¹ 杨环¹ 西永明² 郭建伟² 潘振宽¹ 黄宝香¹

1 青岛大学计算机科学技术学院 青岛 266071

2 青岛大学附属医院脊柱外科 青岛 266000

(ttong_liu@163.com)

摘要 脊柱疾病是现代生活中常见疾病之一,目前其诊断与治疗主要依赖于医生的专业水平和临床经验,这样不仅给医生带来沉重负担,而且效率低下。以神经网络为代表的机器学习算法能够自动提取脊柱数据集中的特征信息,辅助医生快速定位病灶区域,实现精准治疗。文中从实验数据、特征选择、算法模型和性能评估指标等方面,对机器学习技术在脊柱疾病应用中的研究现状进行了系统总结。首先从机器学习算法角度出发,阐述典型算法在疾病诊治中的用途;其次围绕实际应用,从危险因素分析和疾病预测、疾病识别和分类、脊柱图像的特征提取和分割3方面,结合具体实验对比机器学习模型的性能;最后总结目前应用中存在的局限性并提出展望。

关键词:机器学习;神经网络;脊柱疾病;智慧医疗;研究综述

中图分类号 TP399

Review on Intelligent Diagnosis of Spine Disease Based on Machine Learning

LIU Tong-tong¹, YANG Huan¹, XI Yong-ming², GUO Jian-wei², PAN Zhen-kuan¹ and HUANG Bao-xiang¹

1 College of Computer Science and Technology, Qingdao University, Qingdao 266071, China

2 The Affiliated Hospital of Qingdao University Spine Surgery, Qingdao 266000, China

Abstract Spine diseases are prevalent in modern society. The diagnosis and treatment mainly depend on doctors' professional knowledge and clinical experience. More and more patients and conventional treatments resulted in heavy overload and inefficient diagnosis. Machine learning algorithms can automatically extract useful information from datasets and images, assisting doctors to locate the lesion and carry out the accurate treatment. This paper focuses on the applications of machine learning in the field of spine disease and summarizes the relevant research from aspects of datasets, feature selection, model, evaluation metrics, and so on. Firstly, in terms of machine learning algorithms, the utility of some typical algorithms in disease diagnosis and treatment is described. Moreover, in terms of the actual applications of disease diagnosis and treatment (risk factor analysis and disease prediction, disease recognition and classification, feature extraction of spine image and image segmentation), the performances of several important models are compared in some specific experiments. Accuracy, specificity, sensitivity, AUC, and other evaluation indexes are involved. Finally, the major limitations and corresponding issues in current applications are summarized.

Keywords Machine learning, Neural network, Spine disease, Smart health-care, Review

1 引言

近年来脊柱疾病的发病率越来越高,年轻化的趋势越来越严峻^[1],随着工作和生活节奏的加快,手机、电脑的广泛应用,人们的体力劳动减少,静坐时间加长,脊柱疾病已然成为现代社会的常见疾病,常见的脊柱病包括颈椎病、脊柱畸形、脊柱肿瘤、腰椎退行性疾病和腰椎间盘突出症等。

机器学习是计算机科学和统计学的交叉,已经用于疾病辅助诊断、医学图像分类和诊疗决策系统等医学领域。2017年我国发布《新一代人工智能发展规划》,将人工智能上升为国家战略^[2],以机器学习为代表的人工智能技术在医疗领域

的应用呈现快速发展的景象,有些领域(如癌症分类^[3]和心血管病预测^[4])的辅助诊断效果甚至可以媲美专业医生的水平。机器学习技术有助于医师对脊柱数据集的理解和对脊柱影像的识别,很多脊柱外科研究者将其用于辅助诊疗,但目前缺乏标志性的研究成果。本文就机器学习在脊柱疾病中的应用进行综述。以“深度学习/机器学习/人工智能+脊柱”“脊柱+预测/分类/分型”等作为主题词检索中国知网和万方数据库,以算法的英文名+“spine”或“deep learning/machine learning”+“spine”为主题词通过谷歌学术检索外文数据库,检索时间为2000年至今,共检索出文献若干,经过去除内容相似、重复和与主题不相关等文章,最后确定纳入85篇。

基金项目:国家自然科学基金青年项目(61602269);中国博士后科学基金(2017M622136);山东省重点研发计划(公益类专项)(2019GGX101021)

This work was supported by the National Science Foundation for Young Scientists of China (61602269), China Postdoctoral Science Foundation (2017M622136) and Key R & D Program of Shandong Province (2019GGX101021).

通信作者:杨环(cathy_huanyang@hotmail.com)

本文的主要工作内容包括:对监督学习、无监督学习和半监督学习算法在脊柱病辅助诊疗中的应用现状进行总结。其中回归和 XGBoost 算法在预测疾病方面简单高效,贝叶斯算法、决策树和 SVM 常用于疾病分类,聚类和神经网络则常见于图像分割。然后我们从疾病的危险因素分析、疾病识别和分类、脊柱图像特征提取和分割 3 方面对各算法性能进行对比,发现在处理特征数据集方面:回归模型简单迅速,但模型本身无法筛选特征;基于树的模型能够提供属性重要性的度量,在预测分类的同时进行属性选择,该类模型在处理小样本集时泛化能力比较差;对于高维特征的小样本集,SVM 算法的精度和效率皆令人满意。在处理脊柱图像数据集方面:无论是图像的特征提取还是脊柱结构的分割,神经网络均有较高的准确度。最后对当前研究存在的局限性做总结,对未来的发展趋势进行探讨。

2 常用医学机器学习技术

机器学习技术应用于脊柱外科领域,主要有 3 种方式:监督学习、无监督学习和半监督学习方式。本节我们拟对各种机器学习算法的优势与弊端进行介绍。

2.1 有监督学习算法

监督学习是训练机器学习模型最常用的技术。用于脊柱疾病诊治的监督学习算法包括回归算法、贝叶斯算法、决策树、支持向量机、集成学习、神经网络以及最新的深度学习技术,本节对监督学习模型在脊柱疾病诊断领域的应用进行总结。

2.1.1 基于统计学的方法

基于统计学的机器学习算法包括线性回归、逻辑回归和贝叶斯分类算法等。线性回归通常用于解决预测问题,将输入数据表作为多维向量,即一般意义上的特征,利用自变量特征的值预测控制因变量的值,回归系数用来分析特征重要性。逻辑回归是在线性回归的基础上,将输出值限定在 0 到 1 之间,对应疾病发生的可能性,实现二元分类预测^[5]。该类模型可以清晰地表明疾病及其危险因素之间的显著关系,但模型本身无法衡量特征的重要性。逻辑回归模型因其简单易操作,在研究中常用作对比模型。贝叶斯分类算法则是一种利用概率进行分类的算法,它假设特征彼此独立,须联合每个特征的概率计算最终类别的概率,在处理小规模多分类的数据时效果稳定。而实际上医学数据的特征并非彼此独立,因此该模型有时不能直接用于分析医学数据,可以考虑加入先验知识等改进模型。

回归算法可以有效地进行危险因素分析和疾病预测。如:2012 年,Darren 等^[6]利用回归模型评估与椎动脉损伤(Vertebral Artery Injury, VAD)相关的特征重要性,其中显著性特征 AS/DISH 的存在性和横突孔内移位等因素可以预测 VAI 发展的可能性。2014 年,Acaroglu 等^[7]利用多元回归模型分析影响成年脊柱畸形患者生活质量的因素(如:性别和术前术后的 T1 矢状倾斜等),通过对相关因素进行排序,帮助医生更好地理解脊柱侧凸中的问题。2016 年,Xia 等^[8]运用多分类回归模型识别脊柱炎的相关因素,通过对患者的临床特征进行单因素分析,成功实现脊柱炎预测。2018 年,Wang 等^[9]对脊柱侧凸的危险因素进行回归分析,最终将睡眠时间、参加对抗性体育比赛和睡姿 3 个变量纳入模型实现预测。同

年 Pasha 等^[10]基于 X 光图片获取的脊柱参数,利用回归模型预测三维脊柱模型中的指标参数,结果表明多重回归可以准确地预测参数。

利用贝叶斯算法可以有效进行疾病识别和分类。如:2013 年,Peter 等^[11]利用贝叶斯方法检测脊柱 CT 中的骨损伤并对不同大小的损伤进行分类,实验的特异性高于 90%。2014 年,Verma 等^[12]利用贝叶斯分类器对脊髓损伤患者的疼痛进行分级;Alomari 等^[13]提出一种基于贝叶斯理论识别腰椎区疝的方法,在 65 个临床 MRI 上识别疝的平均准确率达到 93.9%。2015 年,Mandal 等^[14]利用贝叶斯网络对脊柱异常的相关特征进行选择,包含骨盆倾斜度和骨盆半径等 6 个属性,最高分类准确率为 96.55%,该方法可以提高脊柱疾病诊断软件的可靠性,增强精确治疗的可行性。

2.1.2 支持向量机

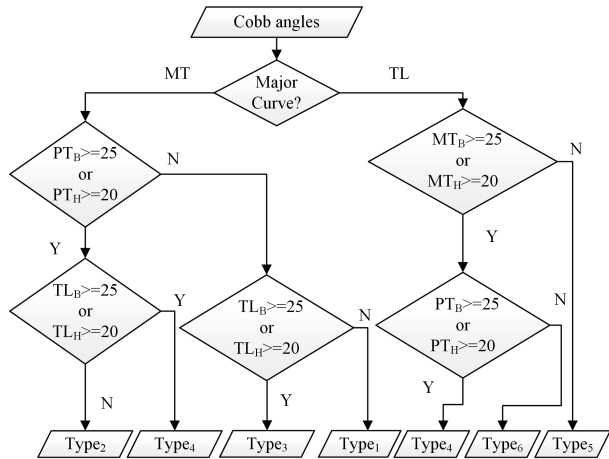
支持向量机(Support Vector Machine, SVM)是基于核函数的分类算法,将输入样本视作多维向量,即多维空间中的点,该算法确定划分各类别点的超平面以区分不同类数据。SVM 算法在脊柱疾病的应用颇多,其优势在于泛化能力强,尤其适合处理高维特征的小数据集。

SVM 算法可以帮助识别脊柱疾病并分类。2007 年,Yao 等^[15]利用计算机辅助系统检测脊柱转移瘤,基于 CT 图像计算 26 个指标特征输入 SVM 训练模型,模型在测试集上的假阳性率为 5.7 个/例。2011 年,他们对该系统进行了改进,加入三维特征的训练,模型的性能得到了提升。2013 年,同组人员^[16]又对胸腰椎棘突硬化疾病进行了研究,将 532 例脊柱硬化性病例分为 10 个对照组进行 SVM 的训练和测试,该方法可以有效识别病灶。2006 年,Ramirez 等^[17]利用 SVM 分类器评估特发性脊柱侧凸的严重程度,基于患者的 12 项指标,测试集的准确率可达到 69%~85%,结果优于线性判别分类器和决策树。2017 年,Jiang 等^[19]利用 SVM 预测腰痛患者的功能康复水平,基于 30 名患者的信息进行实验,与实际结果对比,其分类准确率达 96%以上。2019 年,Tan 等^[18]先利用 Unet 网络对脊柱进行分割,然后利用最小外包络矩形局部逼近椎体的边界,最后利用 SVM 对拟合的脊柱边界曲线进行分型,测试结果与医生测量结果一致。除此之外,SVM 算法也被用于脊柱图像的特征检测和划分,2012 年,Lecron 等^[20]开发了一个包括椎体检测、分割和角度测量的框架进行颈椎 X 光图像分析,利用 SVM 将检测到的兴趣点进行分类,通过角点寻找椎体的边界,最终该方法在 89.8%的病例中成功检测到脊椎。2014 年,Yu 等^[21]基于 SVM 算法对腰椎超声图像进行区域的划分以获取针的插入位点,在测试集的实验准确率为 95%,并在验证集上成功确定了所有正确针插入位点。2014 年,Wang 等^[22]提出将脊柱图像的分割转化为边界回归问题,对临床受试者的脊柱图像进行多维 SVM 回归分割,分割所得的 Dice 相关性系数为 91.2%。2017 年,Wanb 等^[23]结合差异张量成像的自动化框架,使用 SVM 算法将对对照组和颈椎病组精确分类,在 MR 图像的测试准确率为 95.73%,表明该方法可以有效检测病变区域。

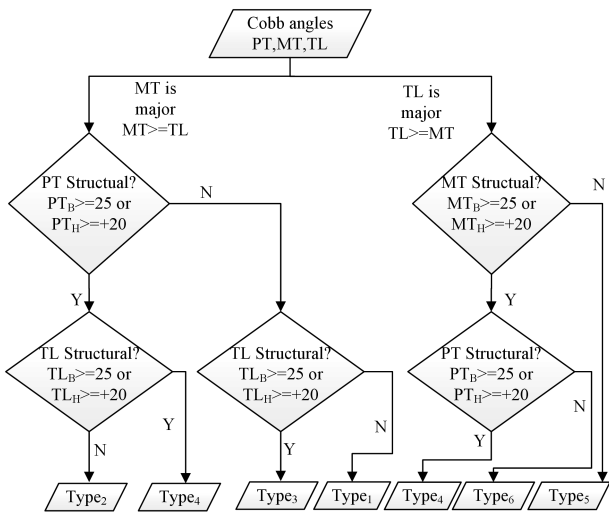
2.1.3 决策树

决策树是一种常见的监督学习的分类算法,利用无规则的数据实例推理有效的分类规则^[24],该方法提供属性重要性的度量,对疾病分类的同时可以进行属性选择,模型的可解释

性强,然而该模型在处理小样本集时的泛化能力比较差。在脊柱疾病的应用中,决策树算法通常用来辅助决策支持系统。例如:2010年,Phan等^[23]利用决策树提高青少年特发性脊柱侧凸分类的准确性,通过比较单独的 Lenke 准则进行分类的准确率和使用决策树之后的准确率,分类精度从 77.2% 提高到 92.9%,分型准则如图 1 所示,数据表明此改善在统计学上是有意义的。2014年,Hanon等^[26]用 3 种策略为 19 岁以下患者的假设人群构建决策分析树,确定儿童钝性创伤中颈椎损伤的最佳初始评估策略。



(a) 结合决策树的 Lenke 分型准则



(b) 用于临床设置的决策树图

注: MT 表示主胸弯; TL 表示胸腰弯/腰弯; PT 表示近端胸弯; XB 表示 Bending 向上 X 弯的 Cobb 角; XH 表示矢状位 X 片上的 Cobb 角

图 1 决策树辅助脊柱侧凸分型诊断^[25]

Fig. 1 Decision tree of automated AIS classifier according to Lenke^[25]

2.1.4 集成算法

集成学习是当下机器学习的热门研究方向之一。集成策略通常包含 Bagging 和 Boosting 两种,其中研究者最常用的 Bagging 模型是随机森林(Random forest, RF),常用的 Boosting 算法有自适应提升算法 Adaboost 和梯度提升算法 XG-Boost。集成方法通过实现模型缺陷的互相弥补,保证机器学习的质量和效率,可以有效地降低疾病的误诊率。

研究表明集成算法在脊柱图像的特征提取方面作用显著。例如,2009年 Huang等^[27]设计了全自动椎体检测和分割系统,利用基于 Adaboost 算法的检测器进行椎体检测,可

实现 98% 的椎体检出率。2012年,Zhan等^[28]采用 Adaboost 级联分类框架和 Haar 小波特征模拟椎体的外观,并利用层次学习和局部定位模型实现 MR 图像中的脊柱检测,椎间盘和椎体的平均检测误差分别为 1.91 mm 和 3.07 mm,说明该方法是有用的。2015年,Smmr等^[29]利用改进的霍夫森林检测颈椎角点,实验误差为 2.08 mm。同年,Robert等^[30]利用 Canny 边缘算子和随机森林回归模型拟合脊椎边界并实现椎骨分割,在 CT 图像上测试的平均 Dice 在 90% 以上。2016年,Cheng等^[31]对 CT 中的脊柱进行定位分割,使用 Adaboost 捕捉局部对比度,利用多个数据集进行实验,平均检测误差为 1.6 mm。2019年,Tao等^[32]利用 RF 和水平集算法分割椎骨,提取图像的特征并利用加权随机森林回归算法获取脊柱中心点位置,测试效果表明随机森林算法可以有效地用于椎体分割。2019年,Ebrahimi等^[33]利用矢状 X 光片的视觉特征进行椎角检测,基于人工标注了脊柱区域和角型的 X 片数据集训练 RF,在受试者图像的评估结果表明该模型可以快速精确地识别椎骨。此外,集成算法可以帮助划分脊柱图像的区域并实现图像分割,2015年,Jebri等^[34]利用颈椎外侧 X 片图像检测退行性变化,每个椎体拟合一个中心点并提取感兴趣的区域,基于区域特征利用随机森林分类器将其划分为退化或正常,交叉验证的准确率超过 95%。2016年,Bromiley等^[35]进行骨质疏松性骨折的检测,利用随机森林回归器定位每个椎体,然后利用随机森林回归投票约束局部模型对每个椎体进行分割,最终 97% 的椎体被分割且平均误差较小。2016年,Lin等^[36]利用 RF 实现椎间盘退行性变的分级,结合图像的灰度信息,算法实现了对前五级退行性变的全自动分级,分级敏感度良好。2016年,Baka等^[37]提出一种基于 RF 算法的自动分割骨骼的方法,在测试集上以 0.82 的精度实现了 0.86 的召回率。2018年,Gawel等^[38]采用级联分类器进行椎骨的自动识别然后利用主动外观模型进行分割,自动分割 MRI 的结果与手动分割结果有很好的 consistency。

XGBoost 算法可以很好地用于疾病预测。2018年,Jia等^[39]基于医院的骨科医疗数据,构建基于 XGBoost 的骨科辅助诊断分类预测模型,结果与决策树和 RF 进行比较,该分类器有更好的预测效果。2019年,Chen等^[40]对特发性脊柱侧凸进展风险预测模型进行了探究,基于患者的个人信息构建模型的输入特征,利用梯度提升决策树进行风险预测,最终验证集的准确率为 90%。同年,Bertsimas等^[41]利用梯度增强树算法进行儿科患者颈椎损伤的预测,测试结果表明该方法是高度准确的,可用来预测儿科颈椎损伤。

2.1.5 神经网络

神经网络是受生物启发的模型,类似于神经元的连接和相互作用,信息通过人工神经元从输入流向输出,人工神经元按层组织,执行简单的操作,如将输入乘以权重进行线性组合,然后通过激活函数处理得到输出,神经网络的训练在于找到组合权重的最佳值,使得最终输出与真实情况最接近^[42]。目前,神经网络已经发展了很多种模型,包括人工神经网络、卷积神经网络和各种深度学习模型。其中应用最多的深度学习是神经网络技术的现代延伸,是具有更多层的神经网络,可以探索数据中更复杂的非线性模式,对未知数据进行预测。在脊柱疾病的相关研究中,神经网络的应用颇为广泛,在脊柱图像分割方面的成就尤为突出。

深度学习的快速发展为图像分割带来了新的解决思路。通过多层变换模仿人脑分析图像,学习逐渐抽象的层次化特征,实现图像语义类别的划分^[43]。2017年,Han等^[44]提出使用递归生成对抗网络 Spine-GAN 分割 MRI 影像,模型通过对 253 例患者的影像进行学习,得到的 Dice 结果为 87.1%。Sekuboyina 等^[45]将 Attention-Net 用于脊柱定位,利用结合数据增强技术的 FCN 进行脊柱分割,模型在 3 个数据集的平均 Dice 值达到 87%,还有提升空间。LPP 等^[46]使用全卷积网络分割第三腰椎横突 CT 图像并测量肌群的横截面积,通过微调 ImageNet 数据集上的预训练模型,最终测试的 Dice 为 93%左右。2018年,Liu 等^[47]利用深度卷积一反卷积网络分割脊柱 CT 图像,利用 MICCAI 数据集训练模型,分割的 Dice 为 89.26%。Li 等^[48]利用 3D 全卷积神经网络分割脊柱 CT,分割的平均 Dice 约为 95%。Zhou 等^[49]设计 N-net 网络用于椎体分割,将残余结构和跳跃连接引入三维全卷积网络,使用 CSI 挑战赛数据进行模型训练,结果的 Dice 值达到 95%左右。Horng 等^[50]利用 U-net, Dense U-net 和 Residual U-net 多种网络分割椎体,使用受试者图像进行实验,结果发现 Residual U-Net 模型的分割效果优于另外两种模型,平均 Dice 值高达 95.1%。2018年,Lessmann 等^[51]基于 5 个不同模式的脊柱数据集,利用全卷积神经网络进行椎体的识别与分割,平均 Dice 值为 94.9%左右。Kim 等^[52]使用边界特定的 U-net 网络(BSU-net)分割椎间盘,应用级联学习方法克服 U-net 池化层的结构局限性,最终分割的 Dice 值比 U-net 高出 3%。Arif 等^[53]使用全卷积网络定位脊柱区域并定位椎体中心,改进 Unet 网络分割椎体,结果 Dice 值为 84%。2019年,Saikumar 等^[54]利用多尺度神经网络工程分割椎体的 MR 图像,Dice 比传统 U-net 模型高出 1.3%。2019年,Liu 等^[55]基于级联卷积神经网络进行脊柱 CT 图像的分割,利用 FCN 进行粗分割,CNN 进行细分割,得到的平均 Dice 值达到 95%,且实验表明模型能够快速实现分节段分割。

神经网络极大地提高了图像特征提取的准确度。例如,2015年,Chen 等^[56]设计 J-CNN 模型识别脊柱 CT 中椎骨的关节,利用 MICCAI 挑战赛数据集评估结果,识别率比决赛的最佳结果高 10.12%。2017年,Hetherington 等^[57]利用卷积神经网络从超声图像中自动识别椎体,20 次超声检测中有 17 次成功识别所有椎体且准确率达 88%。Forsberg 等^[58]基于用图形模型标记椎骨的 MR 图像进行椎体检测,CNN 模型的准确率高于 98.8%。Belharbi 等^[59]利用深卷积网络定位 L3 脊椎,基于 ImageNet 预训练的深层架构利用迁移学习对临床中心数据进行训练,最终验证的平均误差小于 5 mm。2018年,Han 等^[60]设计 DMMLNet 网络解决腰椎处多个器官的定位和分级问题同时诊断腰椎狭窄症,将多输出和多回归学习整合在一个网络中,验证精度高达 84.5%,该方法可用于其他脊柱疾病的诊断。Galbusera 等^[61]利用全卷积神经网络提取患者脊柱模型中的关键点,并根据关键点进行医学特征参数的计算。

神经网络也可以帮助识别脊柱疾病并进行疾病的分类和预测。例如,2003年 Cheruk 等^[62]基于 714 个腰椎 X 光片检测异常椎体,利用多层感知器评估单个椎体的特征,最终识别正常和异常椎体的正确率分别为 90.5%和 86.6%。2014年,Roth 等^[63]通过随机聚集深度卷积网络分类器检测硬化

性脊柱转移,首先利用级联框架生成病变候选区域,然后对候选区域进行随机变换生成二维视图,最后利用 CNN 网络对视图进行分类,最终验证的准确率达 83.4%。之后同组人员^[64]利用深度网络 Convnets 自动检测骨折,通过对 55 个骨折的位置进行学习,检测的准确率达到 85.7%。2015年,Azimi 等^[65]基于患者的临床信息,建立 ANN 模型预测复发性腰椎间盘突出症的程度,最终准确率为 94.1%。之后同组人员^[66]又对腰椎管狭窄症的患者是否需要手术进行预测,通过对 346 名患者的数据进行实验,发现 ANN 的准确率可达 97.8%。Kim 等^[67]利用人工神经网络和 LR 模型预测后腰椎融合的术后并发症,利用 ACS-NSQIP 数据集对模型进行训练和测试,实验发现 ANN 模型预测效果比 LR 的更准确。2017年,Jamaludin 等^[68]利用深度学习模型对腰椎间盘突出进行识别分级,准确率可达 95.6%。同年,Wang 等^[69]利用暹罗神经网络进行脊柱转移检测,病灶检测的准确率为 90%。2018年,Hu 等^[70]使用长短期记忆网络区分健康人群和下腰痛患者,利用 44 名患者和健康个体的运动学特性参数对模型进行训练,最终病人识别的精确度高达 97.2%。Xue 等^[71]基于脊柱 X 光图像进行性别检测,利用已有性别标签的 NHANES 数据集,对微调的 Densenet 模型进行测试,准确率高于 98%。2018年,Bmed 等^[72]利用神经网络预测腰椎间盘患者术后的报告结果,在 422 名患者样本上的预测结果准确率可达 87%,由此可见简单的神经网络即可实现高精度的疾病预测。2019年,Chmelik 等^[73]将计算机辅助诊断系统用于脊柱转移病变的分割和分类,利用卷积神经网络对 277 个病例进行学习,检测病变的平均灵敏度超过 92%。

2.2 无监督学习算法

无监督学习不需要进行人工标注数据,在模式识别和图像处理领域中具有广泛的应用前景^[74]。聚类是最常见的一类无监督学习算法,其他的方法还包括 PCA 降维、马尔可夫随机场、自组织映射神经网络(SOM)等。

2.2.1 聚类

聚类算法通过最大化类内相似度同时最小化类间相似度将相似样本分为一组,输出聚类标签,用来预测分型结果,常用的聚类算法有 K-means 聚类、模糊聚类、层次聚类等。在脊柱疾病研究中,聚类算法可以辅助疾病的识别分类和椎体分割。

在进行疾病的识别与分类方面,Duong 等^[75]利用模糊聚类算法对青少年特发性脊柱侧凸进行分类,通过识别 409 个三维脊柱模型的曲线模式和聚类中心将曲线模式分为 5 类,分型结果与 Lenke 和 King 分类模式类似;2019年,Ames 等^[76]对成人脊柱畸形进行分类并评估手术的质量,基于患者的 SRS22/ODI(健康相关生活质量指标),利用层次聚类探索数据模式,降低术后并发症;Beaudette 等^[77]利用 K-means 识别健康人群的脊柱曲伸模式,结果发现健康人群的模式并不一致,即在分析腰背损伤风险时要具体情况具体分析;Riveros 等^[78]采用 K-means 和 SOM 算法对脊柱患者进行分类,通过对比样本的分类预测结果和实际标签,发现基于 SOM 算法的分类效果更加令人满意,改善了脊椎疾患的检测。此外,模糊聚类是用于脊柱图像分割的常见算法,Athertya 等^[79]和 Ye 等^[80]分别使用模糊聚类算法从 MR 图像中分割椎体,结果均优于阈值分割,且分割效果稳定。

2.2.2 其他

除了聚类算法,文献中用于脊柱外科研究的无监督学习算法还包括主成分分析、马尔可夫随机场和 SOM 等。其中,主成分分析用于对特征数据集或图像中提取的特征降维,SOM 和前面介绍的神经网络模型在结构上类似,包含输入层和竞争层,竞争层用于寻找数据中的规律并进行归类^[81]。这些算法已经用于解决脊柱疾病的许多实际问题。

首先在疾病的识别分类方面,Phan 等^[82]利用 SOM 对成人脊柱侧凸分型进行研究,从大型数据库中提取类似病例指导治疗,是十分有价值的;2018 年,Abdullah 等^[83]利用主成分分析确定对于脊柱异常的最重要的参数,帮助医生快速识别脊柱异常;Zhang 等^[84]提出了一种将 Adaboost 分类器和水平集算法相结合的无监督方法,解决未标记数据的脊柱侧凸识别问题,实验结果表明该方法可以有效地识别脊柱侧弯。其次在术后预测方面,Devries 等^[85]利用无监督学习模型预测损伤患者的术后步行恢复水平,根据数据相似度将 RH-SCIR 数据集聚为行走和非行走两类,最后预测结果与实际结果相似;2010 年,Hao 等^[86]将分割的椎体和局部图像灰度投影结合实现基于马尔可夫随机场的椎间盘快速分割,辅助医生确定腰椎突出面积,治疗腰椎间盘突出等疾病;2016 年,Cai 等^[87]使用变换深卷积网络识别影像中的腰椎,无监督地融合不同模态的图像特征,并自动校正脊椎姿态,在 CT 和 MR 上的识别准确率分别为 96% 和 98%;Patrick 等^[88]利用特征值分解的迭代分割算法识别分割椎体,但精度有待提高。

此外,主动外观模型是图像特征检测的常见算法,通过一组参数控制关键点的位置拟合研究对象的形状,但该方法只单纯利用对象的形状,准确率不高。如:1999 年,Paul 等^[89]使用主动形状模型定位测量图像中可变目标的形状,在 78 个受试者的 DXA 图像上进行椎体特征的检测,AUC 值大于 0.92。2003 年,Xu 等^[90]利用改进的主动外观模型进行脊柱定位,在 X 光图像上的检测效果良好。

2.3 半监督学习算法

半监督学习最大优势在于结合标记数据和未标记数据训练模型。患者数据标记过程需要一些信息的获取,但某些数据的获取可能是有限的,标记所有可用的数据十分困难,采用半监督学习不必为构建分类器而标记所有数据,因此在脊柱疾病的研究中,有学者考虑使用半监督学习算法对数据进行学习。2010 年,Seoud 等^[91]基于采集的部分躯干表面信息预测脊柱侧凸分型,使用有 12 个标签的 90 个训练样本进行学习,效果类似于将 32 个标记数据用于监督学习的准确率,证明了在仅标注少量训练数据的情况下,使用半监督学习算法训练分类器预测脊柱侧弯类型的可行性。2019 年,Kang 等^[92]提出小样本背景下的半监督循环随机森林分类算法,将实际临床数据作为有标签数据,扩增数据作为无标签数据,对 8 种疾病分类效果均优于传统的监督学习算法,在脊柱疾病的应用上值得期待。

3 实验性能对比

我们从疾病危险因素分析与预测、疾病识别与分类、脊柱图像特征提取与分割 3 方面讨论相关算法模型的性能。结合具体实验,从数据集的选择、实验设计和实验结果等方面对模型进行评价,常用的评价指标有准确度、灵敏度、特异度和

ROC 曲线下面积 AUC 等。

3.1 疾病危险因素的分析 and 预测

研究表明疾病的产生及发展过程是非常复杂的,导致疾病的危险因素多种多样,因此对疾病进行预测建模前,首先要进行危险因素的选择。研究中最常用的检验方法是单因素方差分析,基于统计学原理,对连续变量采用 t 检验,对分类变量采用卡方检验或 F 检验,检验 p 值小于 0.05 表示差异有统计学意义,可据此选择显著性特征,尤其是在利用回归算法建模过程中经常用该方法纳入特征变量。此外,一些基于树的模型本身可以度量各特征的重要性来选择危险因素,以上是最常用的分析危险因素的两种方法。确定疾病的危险因素可以更加有针对性地进行预测。例如:骨盆倾斜度和骨盆半径等因素是确定脊柱异常的常见因素,侧弯的 Cobb 角和冠状面躯干失衡参数是对于脊柱侧凸最为显著的因素。本文主要对逻辑回归 LR、K 近邻算法 KNN、神经网络和集成模型等预测算法进行探讨。

(1)神经网络和 LR 模型的对比。Azimi 等^[71]利用人工神经网络预测腰椎间盘突出症。实验数据来自 402 个病例,按 2:1:1 生成训练集、测试集和验证集。基于 SPSS 的方差分析选择特征,将 14 个检验 p 值小于 0.05 的变量作为输入,使用 SPSS 构建 ANN 模型,选择创建多层感知神经网络,设置三层结构:输入层 14 个节点,隐藏层 8 个节点,输出层 2 个节点,分别输出复发或原发突出症的概率,调试学习率和中止条件,将结果与 LR 模型的预测结果进行对比。结果表明,两者的敏感度相仿,神经网络的 AUC 比 LR 高了 7% 左右,ANN 模型具有更好的预测准确率。

(2)集成模型和 LR 模型的对比。2019 年,Bert 等^[41]预测儿科患者的颈椎损伤。实验数据由包含 6 个特征的 9652 名患者样本组成,119 名有颈椎损伤,75% 的随机样本用于训练和测试,其余 25% 的样本作为验证集。建立 3 个预测模型,由决策树集成的 Optimal Trees (OCT) 模型,深度设置为 2,最小叶规模为 25,模型可列出特征的重要性并选择 GCS 总分和年龄两个主要变量画出树的结构,解释性强。使用 Rstudio 建立模型,返回患者损伤的预测概率,通过设置阈值转换为类别预测,阈值选为 0.00439,加权误分类误差最小。该模型将年龄、GCS 总分和运动分量列为最重要的变量,缺乏可解释性。其与 LR 的对比实验结果如表 1 所列,可以看到 XGBoost 模型具有最好的预测精度。

表 1 集成模型和 LR 在预测儿科患者颈椎损伤的性能对比
Table 1 Comparison of ensemble learning and LR in predicting cervical spine injury

	(单位:%)		
Model	LR	XGBoost	OCT
Sensitivity	95.8	94.96	93.28
Specificity	70.63	84.66	82.34
AUC	94.06	96.69	93.28

(3)集成模型和 KNN 模型的对比。Abdullah 等^[85]对脊柱异常进行预测。实验数据集来自 Kaggle 网站,包括 310 个样本,12 个数值型属性,分为正常或脊柱异常两类,按 8:2 和 7:3 两种方式进行训练集和测试集的划分。建立两种模型,一是 KNN 模型,首先基于统计学的 p 值选择危险因素,腰椎滑脱程度的 p 值最小,作为异常检测因素,模型的 K 值设置为 7,该值通过对多次实验效果进行对比得到。二是 RF 模

型,模型的特征重要性排名与统计分析一致,两个模型的实验结果如表 2(a)所列。从实验结果可以看出数据集的划分影响模型的性能,且不同算法对应的最优划分方式不同,对于 RF 模型使用 30% 的测试数据准确率更高,而 KNN 模型表现却相反,总的看来 KNN 的性能优于 RF。Mandal 等^[4]利用不同的数据对脊柱异常进行预测。实验数据集来自患者的 X 光图像,每个样本被描述为 6 个属性,分为正常或脊柱异常两类,80% 的数据作为训练集,其余数据用来测试。使用旋转森林算法(RF+PCA)方法获得属性的重要性排名,并利用多种机器学习算法进行预测,实验结果如表 2(b)所列。从实验得到 Adaboost 的效果优于 RF 和 KNN,且不难发现 PCA 算法可以提高模型的性能。

表 2 脊柱异常预测的性能对比

Table 2 Comparison of models for spine diagnosis

(a) RF 和 KNN 预测脊柱异常				
Model	RF	RF	KNN	KNN
Data	310(8:2)	310(7:3)	310(8:2)	310(7:3)
Sensitivity	84.64%	87.50%	84.52%	86.98%
Specificity	72.73%	78.13%	95.50%	77.00%
Accuracy	82.50%	80.33%	93.25%	88.87%

(b) 多个模型预测脊柱异常				
Model	RF+PCA	KNN	朴素贝叶斯	Adaboost
Data	310	310	310	310
Accuracy	94.82%	93.10%	94.82%	96.55%
Sensitivity	94.8%	93.1%	94.8%	96.6%
Specificity	86.3%	92%	86.3%	92.9%

基于不同数据集,研究人员利用不同的算法实现了多种脊柱疾病的危险因素分析和相关预测,虽然数据来源各不相同,但我们依然可以看出数据量的增加可以提高预测的精度,特征的选择对模型的性能有显著影响。其次与简单的 LR 相比,其他机器学习模型通常具有更复杂的内部逻辑,可以更好地创建自变量和因变量之间的关系,准确度更高。总之,机器学习算法已经越来越多地用于疾病预测,预测的准确性不断提高。

3.2 疾病识别与分类

机器学习是识别疾病类型的有效工具,通过挖掘医学影像的有效信息,获得与疾病高度相关的特征,有效诊断病灶并实现疾病分类,辅助医生提高诊断效率。研究中最常用于疾病识别的模型是卷积神经网络(CNN),用于疾病分类的模型多种多样,本文只对几个典型的分类模型进行分析。

3.2.1 疾病识别

在脊柱病识别方面,我们主要分析 CNN 模型的性能,通过对比相关文献总结该模型用于疾病识别的优势与不足。2014 年,Roth 等^[62]利用 CNN 检测脊柱转移瘤,实验数据为 49 例患者的 CT 图像,共标注 532 个转移瘤的病变区域,通过两层级联框架生成病变候选区域的 2D 视图,然后利用视图训练模型,模型包含多个卷积层、池化层、全连接层,采用交叉验证方式分析模型性能,分析一个 CT 大致需要两分半钟。2017 年,Wang 等^[93]利用暹罗深度神经网络检测脊柱转移瘤,实验数据来自 26 例患者的 MR 图像,模型包括 3 个相同的子网络,对应 3 种分辨率的图像块,每个子网络包括卷积层、归一化层、池化层和全连接层等,子网络设置有相同的参数,模型由 Lasagne 实现,训练集和测试集的比例为 3:1,采用

随机梯度下降方式进行训练, batch size 设为 128, momentum 为 0.9,学习率为 0.06,利用 FROC(free-response ROC)曲线分析检测结果。2019 年,Chmelik 等^[68]利用深度 CNN 检测同种病灶,实验数据由 17 个病例的 MR 切片组成,按 6:2:2 的方式划分训练集、测试集和验证集,模型包含 17 层:卷积层、池化层、rule 激活层、dropout 层和全连接层。利用 FROC 曲线分析性能,每个患者的计算时间为 4 min 左右,3 个实验的部分结果如表 3 所列。

表 3 CNN 识别脊柱转移瘤的实验分析

Table 3 Experiments of metastatic spine cancer via CNN

Time	2014	2017	2019
Data	49CT	26MR	17MR
Sensitivity	80.00%	89.10%	92.00%
Accuracy	83.40%	90.00%	84.00%
FPs	9.5/例	0.2/例	3/例

3 个实验是针对同一种疾病的不同数据集,可以看出神经网络可以有效识别脊柱转移瘤的病灶,随着网络层数的加深,学习的特征越来越丰富,识别准确率越来越高。在构建网络时应注意过拟合问题,3 个模型均有正则化设计层,但明显的局限性是网络训练的参数多,需要大量的训练数据,且训练数据需要有类别标号,通常需要专家来手工标记。

3.2.2 疾病分类

基于各种不同的数据源,结合具体实验本文评估了以下多种机器学习算法进行疾病分类的性能。

(1)SVM 算法。2006 年,Ramirez 等^[17]对脊柱侧凸的严重程度进行划分。实验数据包括 111 名脊柱侧凸患者的 X 光片和临床信息,利用图像计算指标特征计算并进行特征筛选,最终确定纳入 12 项特征,基于 MATLAB 设计 3 种分类器:1)SVM,核函数为径向基函数,使用 LIBSVM 软件包进行建模,针对多分类的情况,采用一对一方式按投票给出最后类别;2)决策树,利用基尼系数进行特征选择,交叉验证判断结果的可靠性;3)线性判别器 LDA。SVM 的实验结果比决策树和 LDA 的高出 7% 左右。但模型本身的分类效果并不理想,因为可用数据量较小,对于小样本集,使用半监督学习可以在一定程度上提高预测精度。如 2010 年,Seoud 等^[91]使用半监督学习对脊柱侧凸曲线进行分型,每个患者包括 45 个特征,半监督学习的效果类似于 32 个标记数据用于监督学习的准确率。此外 RF 泛化能力优于 SVM, Lin 等^[36]利用 RF 对退行性变的分级进行研究,训练专家标注分级的图像,利用 37 例患者的 MR 数据进行测试验证,采用留一法的交叉验证方式,结果可以实现前 5 级的退行性变判定,分类的加权准确率为 91.82%。

(2)贝叶斯算法。Verma 等^[12]对脊髓损伤患者的疼痛进行分级。实验数据来自 48 个患者,分为两个数据集:35 个患者各包含 5 个属性的数据集一和 13 个患者各包含 16 个属性的数据集二。基于贝叶斯理论设计了两种实验算法,假使各属性相互独立的朴素贝叶斯和加入先验知识的贝叶斯网络算法。朴素贝叶斯在两个数据集上的准确率为 63% 和 60%,贝叶斯网络在两个数据集上的准确率为 78% 和 90%。实验结果说明数据集和特征的质量对分类性能影响显著,其次对于医学图像的学习,先验知识至关重要。

(3)决策树算法。Oude 等^[94]进行下腰痛分类的研究。实验的训练集包括 1288 个虚构样本,每个样本包含 15 个属

性,分为3类,测试集来自真实病例,实验设计决策树、RF和Boosting 3种模型,使用RStudio训练模型,测试集上的结果如表4所列。分析可得,增强树模型的平均灵敏度和特异度最高,意味着在真实案例上Boosting模型分类效果最佳。

表4 基于树的模型进行下腰痛分类的实验分析

Table 4 Comparison of models on the classification of low back pain

(单位:%)

Model	决策树	RF	Boosting
Accuracy	71.00	53.00	71.00
Sensitivity	32.00	36.00	37.00
Specificity	68.00	72.00	75.00

(4)无监督学习算法。Riveros等^[80]利用UCI机器学习库的“脊椎数据集”识别脊柱异常患者,采用K-means和SOM两种无监督学习算法,80%的数据进行训练,20%数据用于验证。K-means模型使用MATLAB编写,K设置为3,共进行4次测试。SOM模型使用MATLAB工具箱,共进行1000次实验。实验结果表明SOM的分类效果略优于K-means,其中特异性和敏感度参数均高出0.5%,两者均改善了脊椎异常的检测。

(5)神经网络。Hu等^[66]使用长短期记忆网络区分识别腰痛患者。实验包括44名患者的运动传感器数据和压力中心,训练集包括1073个序列,测试集包括107个序列。网络包含输入层,两个堆叠的LSTM层和softmax层输出正常或异常的概率,模型基于tensorflow框架进行训练,交叉熵函数作为损失函数,学习率设为0.0025,正则项系数设置为0.0015,Batch设为200。利用测试数据集对模型的性能进行评估,最终模型识别病人的精确度高达97.2%。2017年,Bedi等通过分析患者的症状识别腰椎疾病的类型,实验数据中每个病例包含11个变量,实验设计多种分类模型,用Weka进行模型训练,结果如表5所列。可以看到多层感知器进行腰椎病变预测的结果最佳。

表5 多个模型对腰痛患者进行分类的实验分析

Table 5 Comparisons of models on the classification of back pain

(单位:%)

Model	贝叶斯	决策树	RF	SVM	多层感知机
Accuracy	95.71	84.29	97.14	90.00	98.57
ROC	99.90	96.80	100	96.20	100

综合以上实验可以得到:1)算法的性能很大程度上依赖于数据和特征的质量,目前半监督算法可以在一定程度上解决标注数据量少的问题,其次利用PCA等方法获取有效的特征可以提高算法的性能。2)SVM的性能优于线性判别器和决策树,适合处理高维小样本集。3)神经网络模型分类的准确度明显高于其他模型,在数据允许条件下,该算法极具竞争力。

3.3 脊柱图像特征提取和分割

采用机器学习方法可以从X光片、CT、MR和超声图像等医学影像中提取椎体、椎间盘和脊柱形状等信息,事实上定位数据集中的脊柱结构通常是分割脊柱、实现病理特征分类和预测治疗结果全自动的第一步。本节我们将从图像的特征提取和图像分割两部分对相关实验进行分析。

3.3.1 图像特征提取

本节的图像特征主要指图像的几何特征,包括特征点、边

缘和区域特征。在计算机视觉领域,传统的机器学习算法常与特征算子结合检测提取特征,比如特征点描述算子有SIFT,SURF,HOG以及Haar特征描述等,常见边缘检测算子有Roberts,Sobel以及Canny等。该部分我们主要分析SVM、Adaboost和神经网络3种算法的实验性能。

不少学者将SVM算法用于检测点和检测区域的分类。如Lecron等^[20]对颈椎X光片中的兴趣点进行分类。实验数据包包含245个颈椎图像,使用滤波器检测边缘特征,将边缘的交点作为兴趣点,先用SIFT或SURF进行描述,然后基于描述符利用SVM算法进行分类,将兴趣点分为前上角点、前下角点和其他点3类,从而实现椎体的检测。结果表明SURF描述符在检测过程中的性能优于SIFT描述符,准确率为89.8%。2014年,Yu等^[21]对腰椎超声图像的区域进行划分。实验数据包括1440幅图像,首先采用模板匹配法检测关键特征,中线检测法获取沿中线方向的图像特征,然后训练SVM,采用高斯核函数,正则化项系数设为1,训练准确率高达97%,测试准确率为95%。2017年,Wanb等^[23]检测颈椎病变区域。实验数据集为10个不同体素的MR影像,人工标定感兴趣区域,将病变区域的识别看作二分类问题,实验设计3个模型,其中支持向量机和朴素贝叶斯程序来自WEKA,支持张量STM程序(Support Tensor Machine,STM)由设计者编写,3个模型的性能对比如表6所列。可以看出差异张量成像与机器学习相结合是实现区域分类的强有力的策略,在临床上具有潜在的实用价值。

表6 SVM进行图像特征提取的实验分析

Table 6 Experiments of feature extraction based on SVM

(单位:%)

Model	贝叶斯	SVM	STM
Accuracy	80.16	81.52	92.84
Sensitivity	61.29	54.33	84.16
Specificity	87.81	92.43	96.56

Adaboost算法可以用于图像边缘特征的检测。例如,Huang等^[27]开发了自动检测脊椎系统,提出了一种基于Adaboost算法的检测方法。训练数据来自22个脊柱MR数据集,标记了1398个椎体区域,先用最小边界矩形对区域进行裁剪获得样本,然后经过旋转平移等变换进行数据扩增,最后共获取9786个训练样本,基于图像的梯度和强度从样本中提取特征进行模型训练。实验首先训练了线性核和二次核的SVM,测试集的正确率分别为93.7%和94.4%,该方法可以提供合理的划分区域,但不适用于滑动窗口的目标检测,因此采用Adaboost方法,最终椎体检出率为98%,结果说明该方法高效且准确。

深度学习网络的优势在于不需要结合特征描述符而自动提取图像特征。例如,Forsberg等^[58]利用卷积神经网络进行椎体检测,实验数据包括475例腰椎和245例颈椎的MR图像,被随机分成3组,60%用于训练,20%用作测试,剩余的用作验证集,两种数据设置不同的管道检测,网络结构包括卷积层、池化层和全连接层,使用Lasagne库实现模型,结果准确率超过99%,定位误差小于2mm,神经网络的性能远高于其他算法。

通过对该部分实验进行对比分析,我们发现利用传统的机器学习算法提取图像特征的过程实质上是结合视觉领域的

特征描述算子实现分类任务的过程。SVM 算法通过设置合适的核函数解决特征维度高的分类问题; Adaboost 检测算法可以结合滑动窗口检测边缘, 并利用级联结构提高检出率, 可以很好地实现椎体检测; 神经网络模型最大的优势在于精度高, 对噪声不敏感, 但往往需要充足的数据训练大量的参数。

3.3.2 图像分割

图像分割是基于脊柱图像进行疾病诊断的关键步骤, 将图像中的脊柱进行分割并提取相关的指标特征, 为临床诊治提供可靠依据, 辅助医生作出更准确的诊断。已经用于研究的分割算法有基于阈值的算法、主动外观模型和模糊聚类算法等, 近年来最常用的算法是深度学习, 因此本文聚焦于深度学习算法。

医学影像有多种模态, 如 MR、CT、X 光片和超声影像等, 基于不同的图像类型进行实验, 分割模型大多是对两个经典分割网络(U-net 和 FCN)进行改进, 且分割准确度不断提高。例如, Han 等^[44]使用递归生成对抗网络 Spine-GAN 对 MR 图像进行自动分割; Sekuboyina 等^[45]利用结合数据增强技术的 FCN 进行脊柱分割; LPP 等^[46]使用微调的 FCN 模型分割第三腰椎横突 CT 图像; Liu 等^[47]利用深度卷积一反卷积网络实现脊柱 CT 图像的分割; Li 等^[48]利用 FCN 分割脊柱 CT 图像; Zhou 等^[49]设计 N-net 网络用于椎体分割, 将残余结构和跳跃连接引入三维的全卷积网络; Horng 等^[50]利用 Dense U-net 和 Residual U-net 等多种 CNN 网络进行椎体的检测分割; Lessmann 等^[51]利用 FCN 进行椎体的识别与分割; Kim 等^[52]使用边界特定的 U-net 网络从 MR 图像中分割椎间盘, 改进传统 U-net 的卷积和池化层, 通过级联学习方法克服传统 U-net 池化层的结构局限性; Arif 等^[53]使用卷积神经网络定位脊柱区域并实现椎体分割; Saikumar 等^[54]利用多尺度神经网络对脊柱的 MR 图像进行分割; Liu 等^[55]利用级联卷积神经网络对脊柱图像进行分节段分割, FCN 进行粗分割, CNN 进行细分割。最后给出模糊聚类分割应用以作对比, Athertya 等^[81]使用模糊聚类算法从 MR 图像中分割椎体。各分割实验的 DICE 结果如表 7 所列。

表 7 分割算法性能比较

Table 7 Comparison of models for segmentation

Time	Types	Data	模型	DICE/%
2017 ^[44]	MRI	253	Spine-GAN	87.10
2017 ^[45]	CT	多个数据集	2D 和 3D 结合的全卷积网络	92.10
2017 ^[46]	CT	400	FCN	93.00
2018 ^[47]	CT	CSI	卷积-反卷积	89.26
2018 ^[48]	CT	MICCAI	3D 全卷积	95.00
2018 ^[49]	CT	CSI	N-net	95.00
2018 ^[50]	X 片	595	U-net	94.10
			Residual U-net	95.10
			Dense U-net	94.80
2018 ^[51]	CT, MR	多个数据集	三维 FCN+迭代分割	94.90
2018 ^[52]	X 片	124+172	Unet-S	84.00
2018 ^[53]	MR	20 患者	U-net	86.44
			BSU-net	89.44
2019 ^[54]	CT	CSI	FCN+CNN	95.10
2019 ^[55]	MR	42 患者	多尺度神经网络	86.50
2016 ^[81]	MR	16 个	阈值分割	31.67
			k 均值聚类	78.62
			模糊 c 聚类	86.72

注: N-net 表示 N 形三维全卷积网络; Unet-S 的 S 表示使用 Shape-aware 损失函数; BSU-net 指 Boundary specific U-network

分割效果优于其他分割算法, 而且网络模型大多是对 FCN 和 U-net 分割网络的改进, FCN 网络有意融合粗略的语义信息和精细的外观信息, 提高了分割精度, 但训练速度比较慢, 相对而言 U-net 模型是一个轻量级网络, 适合医学图像的分割, 在应用中根据实验数据对模型进行适当的调整可以改善分割的性能。

结束语 机器学习已经用于脊柱病的早期诊断、预测、治疗和疗效评估等各个方面, 而且随着人工智能的不断发展, 以神经网络为代表的机器学习模型得到了有效的训练, 无论是对于文本数据集还是影像数据集的学习, 模型可以从中获取精妙的特征信息, 辅助医生进行高效诊断。但不容否认当前的研究也存在一些问题, 首先针对图像数据集, 基于单模态影像训练模型, 使得模型的鲁棒性比较差, 当然也有学者开始对多模态的数据进行整合分析, 获取更精确的诊断结果, 多模态研究和诊断越来越受到医疗工作者的期盼。其次建立大规模的脊柱病患数据库是决定机器学习模型在该领域成败的关键, 而实际上可用于研究的医学数据量还是很少, 也许接下来机器学习会着重结合小样本的脊柱疾病, 充分发挥迁移学习、半监督学习和增强学习等技术的作用, 促进智慧医疗的进一步发展。相信机器学习技术可以协助医生在影像学等医学领域做出更快、更好的决策。

参考文献

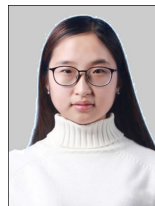
- [1] YUAN L L, ZHANG B Z, DU Q. Analysis of inpatients with Spine disease in children and adolescents[J]. School Health in China, 2013, 34(2): 178-180.
- [2] FENG J F, YU J. Foreword of the frontier progress of artificial intelligence[J]. Journal of Computer Research and Development, 2019, 56(8): 1604-1604.
- [3] YU H L, GU G C, ZHAO J. State of the art on cancer classification problems based on DNA microarray Data[J]. Computer Science, 2010, 37(10): 22-28, 38.
- [4] ZHENG X Y. Research on cardiovascular disease prediction system Based on machine learning[D]. Beijing: Beijing Jiaotong University, 2018.
- [5] RUCZINSKI I, LEBLANC K M. Logic Regression[J]. Journal of Computational and Graphical Statistics, 2003, 12(3): 475-511.
- [6] LEBL D R, BONO C, VELMAHOS G, et al. Vertebral artery injury associated with blunt cervical spine trauma: a multivariate regression analysis[J]. Spine, 2013, 38(16): 1352-1361.
- [7] TAKEMOTO M, BOISSIÈRE L, VITAL J M, et al. Are sagittal spinopelvic radiographic parameters significantly associated with quality of life of adult spinal deformity patients? Multivariate linear regression analyses for pre-operative and short-term post-operative health-related quality of life[J]. European Spine Journal, 2017, 26(8): 2176-2186.
- [8] XIA Q, XU J H, SHUAI Z W, et al. Ordinal Logistic regression analysis of influencing factors of disease activity in ankylosing spondylitis[J]. Acta Universitatis Medicinalis Anhui, 2016, 51(12): 1808-1812.
- [9] WANG H P, SUN Z W, WANG Y H, et al. Prevalence and risk factors of adolescent idiopathic scoliosis in Kunming[J]. Chinese Journal of School Health, 2018, 39(12): 97-100.
- [10] PASHA S, ECKER M, DEENEY V. Considerations in sagittal

通过该部分实验结果的分析不难发现, 深度学习网络的

- evaluation of the scoliotic spine[J]. *European Journal of Orthopaedic Surgery Traumatology*, 2018, 280(2): 1-7.
- [11] PETER R, MALINSKY M, OUREDNICEK P, et al. spine data segmentation and analysis of vertebrae bone lesions[J]. *International Conference of the IEEE Engineering in Medicine and Biology Society*. 2013; 2376-2379.
- [12] VERMA S K, CHUN S, LIU B J. A web-based neurological pain classifier tool utilizing Bayesian decision theory for pain classification in spinal cord injury patients[C]// *Medical Imaging 2014: PACS and Imaging Informatics: Next Generation and Innovations*. International Society for Optics and Photonics, 2014, 9039: 90390E.
- [13] ALOMARI R S, CORSO J J, CHAUDHARY V, et al. Lumbar spine disc herniation diagnosis with a joint shape model[C]// *Computational Methods and Clinical Applications for Spine Imaging*. Springer, 2014: 87-98.
- [14] MANDAL I. Developing new machine learning ensembles for quality spine diagnosis[J]. *Knowledge-Based Systems*, 2015, 730 (Jan.): 298-310.
- [15] WIESE T, BURNS J, YAO J, et al. Computer-aided detection of sclerotic bone metastases in the spine using watershed algorithm and support vector machines[C]// *IEEE International Symposium on Biomedical Imaging: from Nano to Macro*. 2011.
- [16] BURNS J E, YAO J, WIESE T S, et al. Automated detection of sclerotic metastases in the thoracolumbar spine at CT[J]. *Radiology*, 2013, 2680(1): 69-78.
- [17] RAMIREZ L, DURDLE N G, RASOV J, et al. A support vector machines classifier to assess the severity of idiopathic scoliosis from surface topography[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2006, 100(1): 84-91.
- [18] TAN Z Q. Algorithmic study of Lenke classification of idiopathic scoliosis based on U-net[M]. Shenzhen: University of Chinese Academy of Sciences (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), 2019.
- [19] JIANG N, LUK K D, HU Y. A machine learning-based surface electromyography topography evaluation for prognostic prediction of functional restoration rehabilitation in chronic low back pain[J]. *Spine*, 2017, 42(21): 1635-1642.
- [20] LECRON F, BENJELLOUN S, MOHAMMED S. Cervical spine mobility analysis on radiographs; A fully automatic approach [J]. *Computerized Medical Imaging Graphics*, 2012, 360(8): 634-642.
- [21] YU S, TAN K K, SNG B L, et al. Feature extraction and classification for ultrasound images of lumbar spine with support vector machine[C]// *Conf. Proc. IEEE Eng. Med. Biol. Soc.* . 2014: 2014; 4659-4662.
- [22] WANG Z, ZHEN X, TAY K, et al. Regression segmentation for m3 spinal images[J]. *IEEE Transactions on Medical Imaging*, 2015, 340(8): 1640-1648.
- [23] WANG S, HU Y, SHEN Y, et al. Classification of diffusion tensor metrics for the diagnosis of a myelopathic cord using machine learning [J]. *International Journal of Neural Systems*, 2018, 28; 17500364.
- [24] QUINLAN J. Induction on decision tree[J]. *Machine Learning*, 1986, 1.
- [25] PHAN P, MEZGHANI N, MARIE-LYNE N, et al. A decision tree can increase accuracy when assessing curve types according to lenke classification of adolescent idiopathic scoliosis [J]. *Spine*, 2010, 350(10): 1054-1059.
- [26] HANNON M, MANNIX R, DORNEY K, et al. Pediatric cervical spine injury evaluation after blunt trauma: A clinical decision analysis[J]. *Annals of Emergency Medicine*, 2014, 650(3): 239-247.
- [27] HUANG S H, CHU Y H, LAI S H, et al. Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI [J]. *IEEE Transactions on Medical Imaging*, 2009, 280(10): 1595-1605.
- [28] ZHAN Y, MANEESH D, HARDER M, et al. Robust spine detection using hierarchical learning and local articulated model [C]// *Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention- Volume Part I*. 2012.
- [29] ARIF S M M R A, ASAD M, KNAPP K, et al. Cervical vertebral corner detection using haar-like features and modified hough forest[C]// *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*. 2015.
- [30] KOREZ R, IBRAGIMOV B, LIKAR B, et al. Interpolation-based shape-constrained deformable model approach for segmentation of vertebrae from CT spine images[C]// *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer, Cham, 2015: 235-240.
- [31] CHENG E, LIU Y, WIBOWO H, et al. Learning-based spine vertebra localization and segmentation in 3D CT image[C]// *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016: 160-163.
- [32] TAO Y P, JING Y, XU C. IVertebra CT Segmentation Method Based on Random Forest and Chan Vese[J]. *Journal of Chinese Mini-Micro Computer Systems*, 2019, 400(6): 1336-1339.
- [33] EBRAHIMI S, GAJNY L, SKALL W, et al. Vertebral corners detection on sagittal x-rays based on shape modelling, random forest classifiers and dedicated visual features [J]. *Computer Methods in Biomechanics and BIO*, 2019, 70(1/2): 134-146.
- [34] JEBRI B, PHILLIPS M, KNAPP K, et al. Detection of degenerative change in lateral projection cervical spine x-ray images [C]// *Medical Imaging 2015: Computer-Aided Diagnosis: volume 9414*. International Society for Optics and Photonics, 2015: 941404.
- [35] BROMILEY P A, KARIKI E P, ADAMS J E, et al. Fully automatic localisation of vertebrae in CT images using random forest regression voting [C]// *International Workshop on Computational Methods and Clinical Applications for Spine Imaging*. 2016.
- [36] LIN X H. The study of key features extraction and quantization of spinal image in computer-aided clinical diagnosis [M]. Xiamen: Huaqiao University.
- [37] BAKA N, LEENSTRA S, WALSUM T V. Random forest-based bone segmentation in ultrasound[J]. *Ultrasound in Medicine Biology*, 2017, 430(10): 2426-2437.
- [38] DOMINIK G, PAWE G, TOMASZ K, et al. Automatic spine tissue segmentation from mri data based on cascade of boosted classifiers and active appearance model[J]. *Biomed Research International*, 2018, 2018: 1-13.
- [39] JIA W H. Research on Auciliary Diagnosis Model of Orthopedics Based on XGBoost Algorithm[D]. Taiyuan: Taiyuan University of Technology, 2008.
- [40] CHEN M J. Preliminary study on risk prediction model of id-

- iopathic scoliosis[D]. Shanghai: Shanghai University of Sport, 2019.
- [41] BERTSIMAS D, MASIAKOS P T, MYLONASK S, et al. Prediction of cervical spine injury in young pediatric patients; an optimal trees artificial intelligence approach[J]. *Journal of pediatric surgery*, 2019, 54(11): 2353-2357.
- [42] MA R. Principle of Artificial Neural Network[M]. Beijing: China Machine Press, 2010.
- [43] ZHANG R, LI J T. A Survey on Algorithm Research of Scene Parsing Based on Deep Learning[J]. *Journal of Computer Research and Development*, 2020, 57(4): 859-875.
- [44] ZHONGYI H, BENZHENG W, ASHLEY M, et al. Spine-Gan: Semantic segmentation of multiple spinal structures[J]. *Medical Image Analysis*, 2018, 50: 23-35.
- [45] SEKUBOYINA A, KUKAČKA J, KIRSCHKE J S, et al. Attention-driven deep learning for pathological spine segmentation [C]//International Workshop Challenge on Computational Methods Clinical Applications in Musculoskeletal Imaging, 2017.
- [46] LEE H, TROSCHER F M, TAJMIR S, et al. Pixel-level deep segmentation; artificial intelligence quantifies muscle on computed tomography for body morphometric analysis[J]. *Journal of Digital Imaging*, 2017, 30(4): 487-498.
- [47] LIU Z L, CHEN G, SHAN Z Y, et al. Segmentation of Spine ct image based on deep learning[J]. *Computer Applications and Software*, 2018, 35(10): 206-210, 279.
- [48] LI X, HE J. Application of 3D Fully Convolution Network in Spine Segmentation [J]. *Electronic Science and Technology*, 2018, 31(11): 79-83.
- [49] ZHOU W, LIN L, GE G. N-net; 3D fully convolution network based vertebrae segmentation from CT spinal images[J]. *International Journal of Pattern Recognition Artificial Intelligence*, 2019, 33(6): 1957003.
- [50] HORNG M H, KUOK C P, FU M J, et al. Cobb angle measurement of spine from X-ray images using convolutional neural network[J]. *Computational Mathematical Methods in Medicine*, 2019, 2019: 1-18.
- [51] LESSMANN N, VAN GINNEKEN B, DE JONG P A, et al. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification[J]. *Medical Image Analysis*, 2019, 53: 142-155.
- [52] KIM S, BAE W, MASUDA K, et al. Fine-grain segmentation of the intervertebral discs from mr spine images using deep convolutional neural networks; Bsu-net[J]. *Applied Sciences (Basel)*, 2018, 8(9): 1656-1669.
- [53] ARIF S M M R A, KNAPP K, SLABAUGH G. Fully automatic cervical vertebrae segmentation framework for x-ray images[J]. *Computer Methods Programs in Biomedicine*, 2018, 157: 95-111.
- [54] SAIKUMAR K, RAJESH V, RAMYA N, et al. A deep learning process for spine and heart segmentation using pixel-based convolutional networks[J]. *Journal of International Pharmaceutical Research*, 2019, 46(1): 278-282.
- [55] LIU X. Research on CT Image Segmentation algorithm Based on Concatenated Convolutional Neural Network[D]. Harbin: Harbin University of Science and Technology, 2019.
- [56] HAO C, SHEN C, JING Q, et al. Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks [C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, Cham, 2015: 515-522.
- [57] HETHERINGTON J, LESSOWAY V, GUNKA V, et al. Slide: automatic spine level identification system using a deep convolutional neural network[J]. *International Journal of Computer Assisted Radiology Surgery*, 2017, 12(7): 1189-1198.
- [58] FORSBERG D, SJBLOM E, SUNSHINE J L. Detection and labeling of vertebrae in mr images using deep learning with clinical annotations as training data [J]. *Journal of Digital Imaging*, 2017, 30(4): 406-412.
- [59] BELHARBI S, CHATELAIN C, HÉRAULT R, et al. Spotting L3 slice in CT scans using deep convolutional network and transfer learning[J]. *Computers in Biology Medicine*, 2017, 87: 95-103.
- [60] HAN Z, WEI B, LEUNG S, et al. Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning [J]. *Neuroinformatics*, 2018, 16(3/4): 325-337.
- [61] GALBUSERA F, NIEMEYER F, WILKE H J, et al. Fully automated radiological analysis of spinal disorders and deformities; a deep learning approach [J]. *European Spine Journal*, 2019, 28(5): 951-960.
- [62] CHERUKURI M, STANLEY R, LONG R, et al. Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features[J]. *Computerized Medical Imaging Graphics*, 2014, 28(1/2): 99-108.
- [63] ROTH H, YAO, LU L, et al. Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications [C]// Recent Advances in Computational Methods and Clinical Applications for Spine Imaging. Springer, Cham, 2015: 3-12.
- [64] ROTH H R, WANG Y, YAO J, et al. Deep convolutional networks for automated detection of posterior-element fractures on spine CT [C]// Medical Imaging 2016; Computer-Aided Diagnosis. International Society for Optics and Photonics, 2016: 97850P.
- [65] AZIMI P, MOHAMMADI H R, BENZEL E C, et al. Use of artificial neural networks to predict recurrent lumbar disk herniation. *Journal of Spinal Disorders Techniques*, 2015, 28(3): E161-E165.
- [66] AZIMI P, BENZEL E C, SHAHZADI S, et al. Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal canal stenosis; clinical article [J]. *J. Neurosurg Spine*, 2014, 20(3): 300-305.
- [67] KIM J S, MERRILL R K, ARVIND V, et al. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion[J]. *Spine*, 2018, 43(12): 853.
- [68] JAMALUDIN A, LOOTUS M, KADIR T, et al. Issls prize in bioengineering science 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist[J]. *European Spine Journal*, 2017, 26(5): 1374-1383.
- [69] WANG J, FANG Z, LANG N, et al. A multi-resolution approach for spinal metastasis detection using deep siamese neural networks[J]. *Computers in Biology Medicine*, 2017, 84: 137-146.

- [70] HU B, KIM C, NING X, et al. Using a deep learning network to recognize low back pain in static standing[J]. *Ergonomics*, 2018, 61(10):1374-1381.
- [71] XUE Z, RAJARAMAN S, LONG R, et al. Gender detection from spine x-ray images using deep learning[C]// 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS). IEEE, 2018; 54-58.
- [72] STAARTJES V, DE WISPELAERE M, VANDERTOP W, et al. Deep learning-based preoperative predictive analytics for patient-reported outcomes following lumbar discectomy; Feasibility of center-specific modelling[J]. *The Spine Journal*, 2019, 19(5):853-861.
- [73] CHMELIK J, JAKUBICEK R, JAN J, et al. Fully automatic CAD system for segmentation and classification of spinal metastatic lesions in CT data[C]// World Congress on Medical Physics and Biomedical Engineering 2018. 2019.
- [74] SONG Y Q, XIE C H, ZHU Y Q, et al. Research on Medical Image Clustering Based on Approximate Density Function[J]. *Journal of Computer Research and Development*, 2006, 2006(11): 1947-1952.
- [75] DUONG L, CHERIET F, LABELLE H. Three-dimensional classification of spinal deformities using fuzzy clustering[J]. *Spine*, 2006, 31(8):923-930.
- [76] AMES C P, SMITH J S, PELLISÉ F, et al. Artificial intelligence based hierarchical clustering of patient types and intervention categories in adult spinal deformity surgery; towards a new classification scheme that predicts quality and value[J]. *Spine*, 2019, 44(13):915-926.
- [77] BEAUDETTE S M, ZWAMBAG D P, GRAHAM R B, et al. Discriminating spatiotemporal movement strategies during spine flexion-extension in healthy individuals[J]. *The Spine Journal*, 2019, 19(7):1264-1275.
- [78] RIVEROS N A M, ESPITIA B A C, PICO L E A. Comparison between K-means and self-organizing maps algorithms used for diagnosis spinal column patients[J]. *Informatics in Medicine Unlocked*, 2019, 16:100206.
- [79] ATHERTYA J, KUMAR G S. Fuzzy clustering based segmentation of vertebrae in T1-weighted spinal MR images[J]. *Computer Vision and Pattern Recognition*, 2016, 60(2):23-34.
- [80] YE W, TAO J, CHEN X Y, et al. Study of A Novel Automatic Segmentation Algorithm for MR Spine Image[J]. *China Medical Devices*, 2018, 33(9):61-64.
- [81] WAGNER C S, LEYDESDORFF L. Network structure, self-organization, and the growth of international collaboration in science[J]. *Research Policy*, 2005, 34(10):1608-1618.
- [82] PHAN P, MEZGHANI N, WAI E K, et al. Artificial neural networks assessing adolescent idiopathic scoliosis; comparison with Lenke classification[J]. *Spine Journal*, 2013, 130(11):1527-1533.
- [83] ABDULLAH A A, YAAKOB A, IBRAHIM Z. Prediction of spinal abnormalities using machine learning techniques[C]// 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA). IEEE, 2018; 1-6.
- [84] ZHANG L, ZHAO J, YANG H, et al. Unsupervised scoliosis diagnosis via a joint recognition method with multifeature descriptors and centroids extraction[J]. *Computational Mathematical Methods in Medicine*, 2018, 2018:1-14.
- [85] DEVRIES Z, HODA M, RIVERS C S, et al. Development of an unsupervised machine learning algorithm for the prognostication of walking ability in spinal cord injury patients[J]. *The Spine Journal*, 2020, 200(2):213-224.
- [86] HAO S J, ZHAN S, JIANG J G. Spine Disc MR Image Analysis Using Improved Independent Component Analysis Based Active Appearance Model and Markov Random Field[J]. *Journal of Biomedical Engineering*, 2010, 10(27):12-15, 21.
- [87] CAI Y, LANDIS M, LAIDLEY D T, et al. Multi-modal vertebrae recognition using transformed deep convolution network[J]. *Computerized Medical Imaging Graphics*, 2016, 51:11-19.
- [88] PATRICK J, INDU M G. A semi-automated technique for vertebrae detection and segmentation from CT images of spine[C]// 2016 International Conference on Communication Systems and Networks (ComNet). 2016.
- [89] SMYTH P P, TAYLOR C J, ADAMS J E. Vertebral shape; automatic measurement with active shape models[J]. *Radiology*, 1999, 211(2):571-578.
- [90] TAO X U, CAI Y X. Localization of object (spine) in medical image using active shape models[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2003, 2003(2):211-217.
- [91] SEOUD L, ADANKON M M, LABELLE H, et al. Towards non invasive diagnosis of scoliosis using semi-supervised learning approach[C]// Image Analysis Recognition, International Conference. Iciar, Póvoa De Varzim, Portugal, June, Part II, 2010.
- [92] KANG Y. Research and Application of Auxiliary Diagnosis Method Based on Amplified Collaborative Classification of Small Sample Clinical Data[D]. Xi'an: Xidian University, 2019.
- [93] WANG S H, TANG C, SUN J. Multiple sclerosis identification by 14-layer convolutional neural network with batch normalization, dropout, and stochastic pooling[J]. *Frontiers in Neuroscience*, 2018, 12:818.
- [94] NIJEWEEME-D' HOLLOSY W O, VAN VELSEN L, POEL M. Evaluation of three machine learning models for self-referral decision support on low back pain in primary care[J]. *International Journal of Medical Informatics*, 2018, 110(FEB.):31-41.
- [95] BEDI R, SHARMA S. Classification algorithms for prediction of lumbar spine pathologies[C]// International Conference on Advanced Informatics for Computing Research. 2017; 42-50.



LIU Tong-tong, born in 1997, postgraduate. Her main research interests include machine learning and medical data mining.



YANG Huan, born in 1985, Ph.D, associate professor, postgraduate supervisor, is a member of China Computer Federation. Her main research interests include image/video processing and image quality assessment.