

手骨 X 光片骨龄预测中图像预处理的研究

苏 叶, 李 婧, 徐寅林

(南京师范大学计算机与电子信息学院, 江苏 南京 210023)

[摘要] 骨龄预测时,手骨 X 光片经常存在标尺、伪影、噪声、曝光不当等缺陷. 采用常规的滤波加深度学习神经网络等模型进行预测往往正确率不高. 提出一种 X 光片骨龄辅助预测的预处理方法,包括使用专门用于生物医学图像分割的 U-Net 网络将 X 光片中手骨分割出来,使用图像二值化对 U-Net 生成的掩模进行去除背景处理,使用灰度直方图均衡的办法解决图像过亮或过暗的问题. 经上述精细预处理后,再进行深度学习神经网络预测,实验结果表明精细的预处理对实验结果有很好的改进作用.

[关键词] 骨龄预测,预处理,灰度直方图均衡,U-Net 网络,深度学习

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2021)02-0054-06

Research on Image Preprocessing in Predicting the Bone Age of Hand Bone X-ray Films

Su Ye, Li Jing, Xu Yinlin

(School of Computer and Electronic Information, Nanjing Normal University, Nanjing 210023, China)

Abstract: Problems of scale, artifact and noise in the X-ray pictures of hand bone always exist during the bone age prediction. The accuracy of prediction by normal filter and end-to-end deep learning neural network model is not so high normally. This paper uses the method of U-Net, which is used for biomedical image segmentation, to split hand bone X-ray film. The method uses image binarization to remove the background of mask generated by U-Net, and uses grayscale straight square equalization to solve the problem of brightness of images. Deep learning neural network prediction after fine pretreatments mentioned above can significantly improve the results of bone age prediction.

Key words: bone age prediction, pretreatment, grayscale histogram equalization, U-Net network, deep learning

骨龄(bone age)是骨骼年龄的简称,能够反映被测试者的生长发育水平和成熟程度. 骨龄不仅可用以判断个体的生物年龄,还可用于预测儿童成年后的身高. 骨龄预测在临床中常常被用于儿童内分泌和生长发育等相关疾病的诊断,在青少年体育赛事中常被用以确定运动员的实际年龄,以判定参赛资格. 骨龄作为预测生物年龄最权威的评价指标,被广泛应用于临床医学^[1]、体育竞技^[2]和司法判案^[3]等领域,因此,骨龄的准确检测非常重要. 在传统临床领域,医生通过观察非惯用手(一般为左手)手腕部位的 X 光图像得出个体的骨龄,其中观察区域包括左手掌指骨、腕骨及桡尺骨下端的骨化中心的发育程度,手骨 X 光图像及骨关节感兴趣区(region of interest, ROI)如图 1 所示.

传统的骨龄识别方法中,常用的有 GP 图谱法^[4]、TW2 计分法^[5]、TW3 计分法^[6]等. 这些方法都有不同的量化指标和鉴定方法,虽然都可实现骨龄预测,但均存在识别所需时间较长、受医生主观因素影响较大和随机性较大的问题. 如何解决这些问题,使得骨龄预测自动化,提高骨龄预测的准确率,已成为国内外研究人员一直以来的研究目标. 自 20 世纪 80 年代开始,研究人员开始在骨龄预测研究中尝试使用计算机视觉等相关技术,这些方法可有效地解决所需时间长、随机性大、可重复差等问题,从而实现自动化的骨龄辅助预测^[7-13]. 但这些算法只是在作者自己采集的质量较高的图片集上进行验证,具有较大的局限性,或是实验结果不够理想. 近年来,也有研究人员在公开的数据集基础上进行了骨龄预测的算法研究^[14-15],这些方法虽相较于前者更具通用性,且正确率也有了一定的提高,但在其评价正确率指标时,剔除了数据

收稿日期:2020-09-16.

通讯作者:徐寅林,博士,教授,研究方向:精密仪器设计. E-mail: xuyinlin@njnu.edu.cn

集中质量差的图像,因此并未从根本上解决较差质量图像的处理问题。

由于骨龄的公开数据集较少,本文研究对象与文献[14-15]相同,即2017年儿科骨龄挑战赛 RSNA 的公开数据集。该数据集包含12 611个训练图像,其中既有大量优质图片,也包含了许多劣质的图片。常规图像的预处理结果如图2所示。如图2(a)所示,这是由于放射图像在比例、方向、曝光等方面各不相同,通常还存在标尺、伪影、噪声等问题,因而该数据集是考验算法鲁棒性的极佳对象。数据集存在的这些问题,对骨龄预测结果的正确性影响较大,虽然通过常规的图像预处理可使图片质量得到部分改善,但反映骨龄变化的骨骼特征既有细微的边沿形态特征,也有因钙化层次不同的明暗特征,常规的图像处理往往无法兼顾这些细节和区域的特性,因而不能达到理想的预处理效果。如图2(b)所示使用中值滤波法处理图片,虽然去除了图像上的标记,但手骨区域却变得模糊。如图2(c)所示使用图像增强算法,虽然手骨区域对比度增强,但标尺边框部分也一并增强。本文通过 U-Net 模型将手骨图像经分割去除背景、掩模图像二值化、局部灰度直方图均衡等一系列预处理,使得图像在相同的神经网络骨龄预测模型中预测结果得到改善。



图1 手骨 X 光图像及骨关节 ROI

Fig. 1 X-ray image and ROI of hand bone

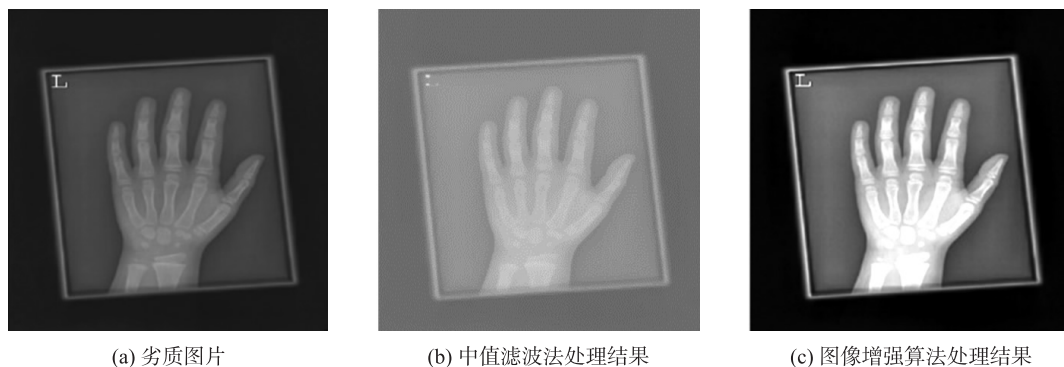


图2 常规的图像预处理结果

Fig. 2 Results of regular image

1 图像预处理

本文使用的预处理系统如图3所示,包括使用 U-Net 去除图像背景等无用信息、进行掩模图像二值化处理以及图像局部灰度直方均衡化等步骤。图像分割处理时,首先将数据集图像统一大小,接着使用分批-迭代的方式通过 U-Net 网络得到所有图像的掩模,最后将掩模与原图融合得到所有分割后的图像。图像增强时,进一步将去除背景后的图像进行局部灰度直方图均衡处理,最终使得所有图像大小统一、去除背景、明暗分布相同。

1.1 基于 U-Net 网络的完整手骨分割

原始数据集图像背景中包含标尺、伪影、噪声等信息,为了消除这些多余信息的干扰,突出与骨龄有关的主体部分,本文使用 U-Net 分割图像去除背景。

1.1.1 图像对应掩模(Mask)的获取

图像对应掩模的获取主要由以下3个步骤组成:

(1)由于 U-Net 网络模型要求输入的图片大小统一,故第一步是将数据集中大小不一的图片都统一为 512×512。

(2)取数据集中 100 张手骨 X 光图像,使用 Labelme 工具手动标注手部掩模,确保有足够的标签能够

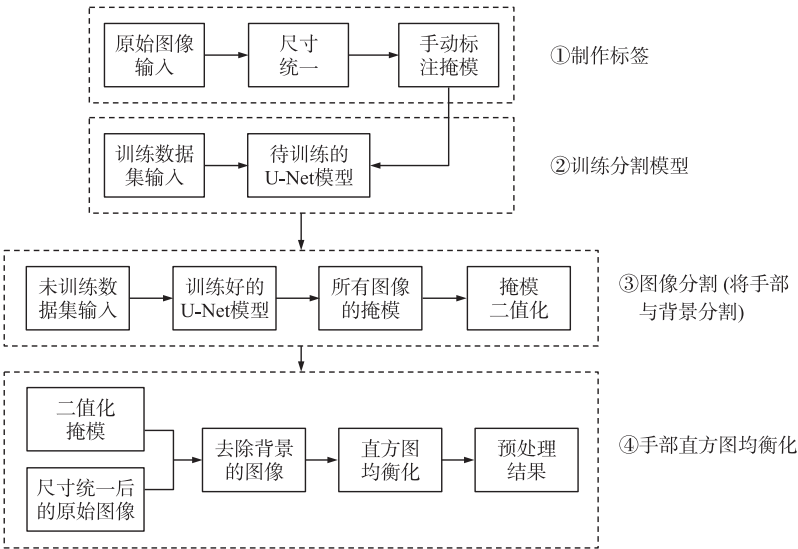


图 3 图像预处理流程图

Fig. 3 Flowchart of image preprocessing

训练得到较好的 U-Net 模型,以便进一步把手骨部分从图像中分割出来.

(3)使用(2)中手动标注完含有标签的图像作为 U-Net 网络的训练集,开展 U-Net 网络的第一次训练,并使用训练出来的模型分割第一个测试集. 由于手动标注的掩模训练集数量较少,且训练集中没有均衡的包含标尺、伪影、噪声、曝光等情况的样本,无法训练出优异的模型,所以在测试时其分割效果参差不齐.

由于数据集容量较大,为了能够得到完美的分割结果,本文采用分批-迭代的训练方式,流程图如图 4 所示. 将数据集以 200 张为单位进行分组,第一次训练后使用分割结果精准的图片直接作为训练集的扩充,其余的图片考虑到由于是第一批手动标注的图片中所含此类特征样本较少而导致无法准确分割,所以进一步将这些图片进行手动标注再作为训练集的扩充进行第二次训练,第二次训练后只有少数图像没有得到完美的手部掩模. 考虑到训练样本过多可能会导致过拟合现象,在第三次训练后,只将效果不好的图片手动标注后扩充训练集. 经过 5 次迭代训练,最终得到训练集样本 631 张,训练得到最佳模型的损失函数 loss 值稳定在 99.89,并成功预测了剩余的图片.

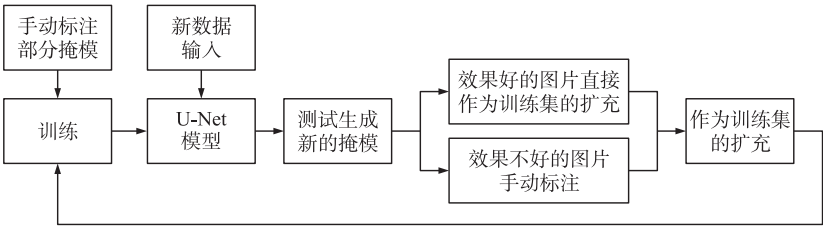


图 4 分批-迭代的 U-Net 训练

Fig. 4 Batch-iterative U-Net training

1.1.2 图像与对应掩模(Mask)的融合

得到图像的所有掩模后,应当使掩模与原图融合从而将手骨区域分割出来. 将图像与上述步骤得到的掩模像素值融合,得到图 5 所示结果. 有些图像手骨区域分割效果较好,但背景部分仍有部分阴影没有

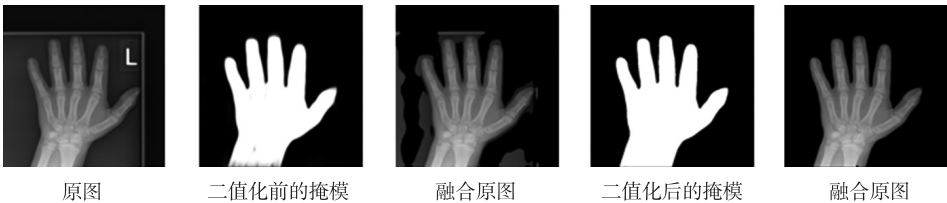


图 5 掩模二值化前后对比

Fig. 5 Comparison before and after mask binarization

完全屏蔽. 分析原因发现 U-Net 网络输出的掩模并非是完全二值化图像,因此需要对掩模做二值化处理. 二值化后的掩模与原图融合后得到完美的分割效果,所有图片都能完整地将手骨区域分割出来.

1.2 手骨 X 光图像局部直方图均衡化

尽管通过 U-Net 分割等处理能够完全滤除图像的背景噪声,但由于拍摄 X 光片时医疗条件、环境和设备等差异,导致拍摄的 X 光图像存在过曝光或欠曝光的情况,这直接使得图像存在明显的明暗差异. 图 6 所示为两张明暗差异较大的图像及其对应的灰度直方图. 实验计算表明,明暗差异较大的图像会对模型的训练及预测产生影响,所以需进一步对去除背景后的手部图像进行局部直方图均衡化的工作.

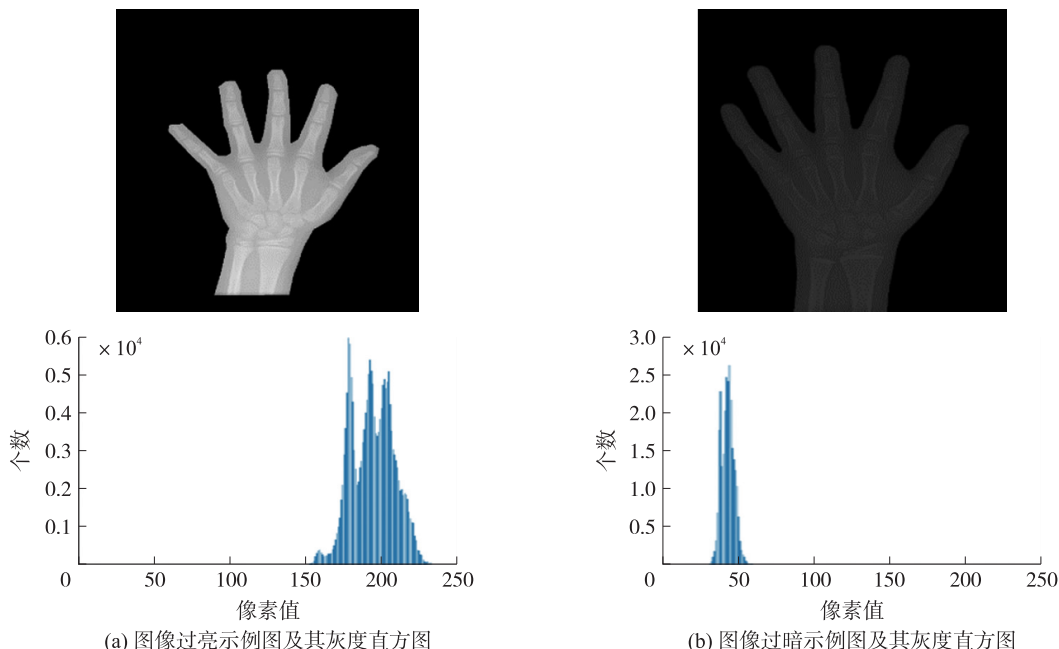


图 6 明暗差异较大的图像手部亮度及其灰度直方图差异

Fig. 6 The differences of brightness and grayscale histogram in the part of hand in pictures

直方图均衡化是通过图像直方图来增加图像的全局对比度,使有用数据可以在直方图上更好地分布,从而在图像中凸显出有用像素,使对比度更加明显. 直方图均衡化对于明暗差异不明显等对比度不高的图像很有效果,尤其是可以使 X 光图像中骨骼结构更加明显.

在直方图均衡化时,通常对直方图先进行归一化. 设灰度图像的维数是 $M \times N$, 像素总数为 MN , 则归一化直方图表示为:

$$p(r_k) = n_k / MN, \quad k = 0, 1, \dots, L-1, \quad (1)$$

式中, $p(r_k)$ 为灰度级 r_k 在图像中出现的概率,且所有分量之和为 1; 假设 r 为待处理图像的灰度,取值范围为 $[0, L-1]$, 则 $r=0$ 表示黑色, $r=L-1$ 表示白色. 直方图均衡化的过程对应于一个变换 T :

$$s = T(r), \quad 0 \leq r \leq L-1, \quad (2)$$

则原灰度级为 r_k 的像素点在均衡化后的灰度级 S_k 为:

$$S_k = T(r_k) = (L-1) \sum_{j=0}^k P_r(r_j) = \frac{L-1}{MN} \sum_{j=0}^k n_j, \quad k = 0, 1, 2, \dots, L-1, \quad (3)$$

式中, MN 为图像像素总数; n_k 表示灰度为 r_k 的像素个数; L 是图像可能的灰度级数量.

对手部使用部分直方图均衡化操作前后的手骨 X 光图像及其灰度直方图如图 7 所示. 可以看出, 图像前景的对比度被明显增强, 所需要的手骨感兴趣区域中的有用信息能够更直观更清晰地展示.

2 基于深度学习的骨龄预测

为了考察上述 X 光片预处理结果对骨龄辅助预测的影响, 本文采用端到端的深度学习神经网络模型进行预测, 本文的结构模型如图 8 所示.

考虑到骨骼的成熟程度因性别而异(男女骨骼发育程度差异最长可达 2 年), 若忽略这种根本差异将

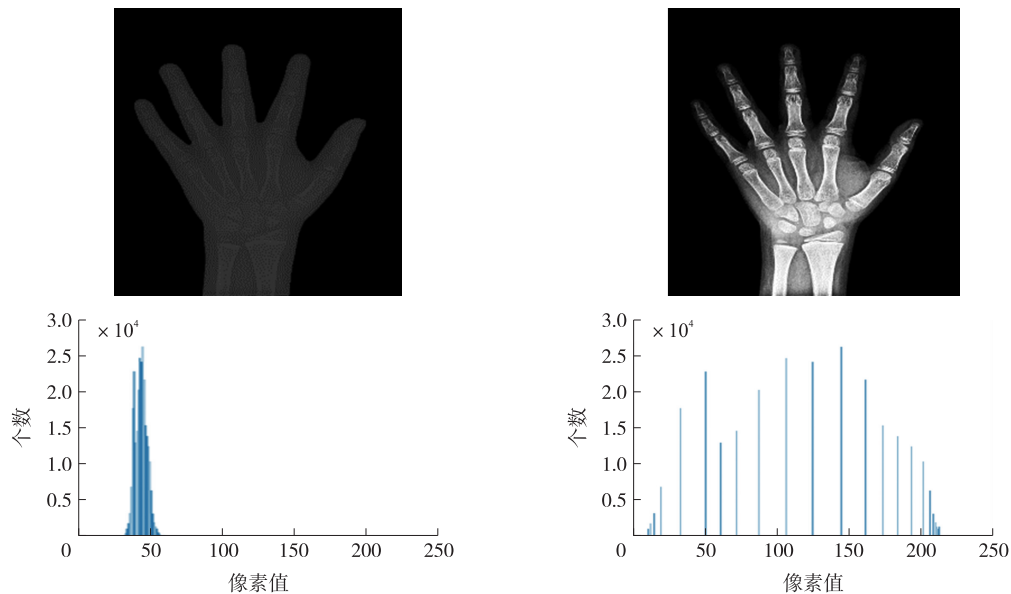


图 7 局部直方图均衡化前后对比(手部)
Fig. 7 Histogram equalization comparison(the part of hand)

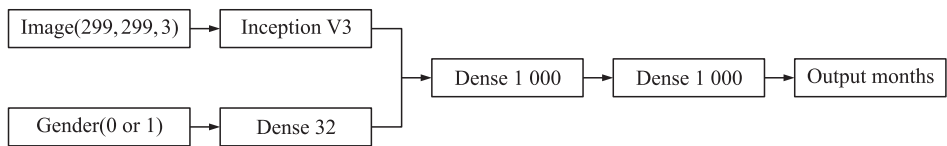


图 8 网络结构模型
Fig. 8 Network structure model

损害网络的性能,故本文中的结构模型将性别信息与图像信息进行了级联。

该模型将图像与性别作为两路单独的输入. 图像信息通过 Inception V3 提取特征,性别的输入为二进制性别信息(女性为 0,男性为 1). 图像输入特征提取后,取最后一个连接层的层铺平,然后与通过 32 层全连接层的性别信息进行级联,接入连接层. 随后通过“relu”激活层将级联后的连接层通过两个 1 000 层密集连接的全连接层,将前面提取的特征联合起来.

该设计思路考虑到每个输入(像素和性别)的相对贡献. 在该模型的串联层,像素贡献了 100 384 个输入,而性别贡献了 32 个. 选择此比率,既不会使性别输入对网络产生过分偏见,又可以赋予其影响整体预测的能力. 额外的密集层为网络提供了更多可学习的参数,可在训练期间进行调整,以使其推断出像素与性别信息之间的关系.

3 实验与分析

3.1 预测模型训练

本文图像预处理基于 PyCharm 2020.1.3 实现,Linux 操作系统上搭建神经网络模型,具体的实验环境如表 1 所示.

本文从 RSNA 数据集所得的 12 611 个训练图像中选出 200 张包含各年龄段、性别分布均匀的图片作为测试集. 训练数据包含 5 778 名女性和 6 833 名男性的放射图像,年龄从 1 个月到 228 个月不等,受试者多为 5-15 岁的儿童.

3.2 预处理结果对照分析

将实验数据按 85:15 的比例分为训练集和验证集,得到 10 720 个训练图像和 1 892 个验证图像. 将网络模型分别在原图数据集 (ID-RD+)、原图与掩模融合的数据集 (ID-MF) 及掩模融合去除背景后进行手部灰度直方均衡处理的数据集 (ID-MF-HE) 中训练 500 个

表 1 具体实验环境	
Table 1 Specific experimental environment	
名称	指标
操作系统	Linux
GPU	NVIDIA-2080 Ti * 2
内存	22GB
Cuda 版本	Cuda = 10.0.130
Keras 版本	Keras = 2.1.6
TensorFlow 版本	TensorFlow = 1.13.1

Epoch. 将本文训练的模型结果与文献[14-15]进行对比(其中, 文献[14]边增亚方法以 BZY 表示, 文献[15]林珏伟方法以 LJW 表示), 实验结果如表 2 所示, 实验结果以平均绝对误差(MAE)作为评价指标, 其公式为:

$$\text{MAE}(X, h) = \frac{1}{m} \sum_{i=1}^m |h(x_i) - y_i|. \quad (4)$$

由实验对比结果可知, 使用 U-Net 去除背景后, 预测误差下降较多, 说明深度学习方法受原始图片背景中无关信息影响较大; 经手部直方图均衡处理后, 误差进一步下降, 但下降幅度较小, 可知图像明暗程度对深度学习方法影响较小. 可知, 有效的精细预处理对骨龄预测结果的精确度有一定的好处. 本文进一步在原图模型及使用了预处理的模型中用 100 张图片进行自动化的预测, 预测时间对比如表 3 所示, 可知预处理未过多增加预测时间.

表 2 对比实验结果

Table 2 Comparative experiment results			
预测方法	MAE/月	预测方法	MAE/月
BZY	7.91	ID-MF	6.15
LJW	9.50	ID-MF-HE	5.44
ID-RD	8.88		

表 3 预测时间对比实验结果

Table 3 Comparative experiment results of prediction time	
实验类别	100 张图片预测时间/s
原图	37
包含预处理	54

4 结论

本文针对 RSNA 提供的用于骨龄预测的手骨 X 光片数据集存在的标尺、伪影、噪声等问题, 提出一种基于精细的预处理和深度学习下的 X 光片骨龄辅助预测模型, 使用了包括 U-Net 分割模型、图像二值化、灰度直方均衡等方法对 X 光片进行预处理, 并使用基于 DenseNet 和 InceptionV3 的端到端网络对数据集进行训练和测试. 实验结果表明, 本文提出的预处理能够有效提升骨龄预测结果的准确性.

[参考文献] (References)

- [1] WILSON D M. Regular monitoring of bone age is not useful in children treated with growth hormone[J]. Pediatrics, 1999, 104 (Suppl 5): 1036-1039.
- [2] 黄陈力子. 中国儿童青少年乒乓球分龄赛运动员骨龄检测状况的研究[D]. 北京: 北京体育大学, 2015.
- [3] 陆慧玲, 郑晶, 顾学安, 等. 青少年活体年龄推断的方法比较[J]. 中国法医学杂志, 2002, 17(增刊 1): 27-29.
- [4] GREULICH W W, PYLE S I. Radiographic atlas of skeletal development of the hand and wrist[J]. The American Journal of the Medical Sciences, 1959, 238(3): 393.
- [5] TANNER J M, WHITEHOUSE R H, CAMERON N, et al. Assessment of skeletal maturity and prediction of adult height(TW 2 method)[M]. London: WB Saunders, 2001: 1-110.
- [6] TANNER J M, HEALY M R J, GOLDSTEIN H, et al. Assessment of skeletal maturity and prediction of adult height(TW 3)[M]. London: WB Saunders, 2001: 243-254.
- [7] KASHIF M, JONAS S, HAAK D, et al. Bone age assessment meets SIFT[C]//Medical Imaging 2015: Computer-Aided Diagnosis. Orlando, USA: International Society for Optics and Photonics, 2015.
- [8] 王燚, 钟俊航. 基于手部 X 光图像的骨龄检测系统设计与实现[J]. 信息系统工程, 2015, 32(1): 28-30.
- [9] 李新华, 赵娟, 袁振宇, 等. 基于 k-余弦曲率和 WSVM 的骨龄识别方法[J]. 计算机应用与软件, 2015, 32(8): 158-161.
- [10] 刘洁琳. 计算机辅助桡骨骨龄等级评估[D]. 北京: 北京交通大学, 2017.
- [11] SPAMPINATO C, PALAZZO S, GIORDANO D, et al. Deep learning for automated skeletal bone age assessment in X-ray images[J]. Medical Image Analysis, 2017, 36: 41-51.
- [12] LEE H, TAJMIR S, LEE J, et al. Fully automated deep learning system for bone age assessment[J]. Journal of Digital Imaging, 2017, 30(4): 427-441.
- [13] MUTASA S, CHANG P D, RUZAL S C, et al. MABAL: a novel deep-learning architecture for machine-assisted bone age labeling[J]. Journal of Digital Imaging, 2018, 31(4): 513-519.
- [14] 边增亚. 基于深度卷积神经网络的自动骨龄识别方法研究[D]. 北京: 北京交通大学, 2019.
- [15] 林珏伟. 基于深度学习的骨龄评估方法研究[D]. 杭州: 浙江工业大学, 2019.

[责任编辑: 严海琳]