

❖ 骨骼肌肉影像学

Automatic measurement model for recognizing key points and automatic measurement of key angles of lower limb alignment on anteroposterior full-length lower limb radiograms based on deep learning

ZHANG Zijian^{1,2}, MA Jianxiong^{1,2*}, BAI Haohao^{1,2}, WANG Ying^{1,2}, SUN Lei^{1,2},
LU Bin^{1,2}, GAO Zhihao^{1,2}, YU Chengshuang^{1,2}, SUN Hanchen^{1,2}, MA Xinlong^{1,2}

(1. Institute of Orthopedics, Tianjin Hospital, Tianjin University, Tianjin 300050, China;

2. Tianjin Key Laboratory of Orthopaedic Biomechanics and Medical Engineering,

Tianjin 300050, China)

[Abstract] **Objective** To construct an automatic measurement model of key angles on anteroposterior full-length lower limb radiograms based on deep learning (DL), and to evaluate its clinical application value. **Methods** Totally 634 anteroposterior full-length lower limb radiograms were retrospectively selected, and the key points of lower limb alignment were marked by 5 orthopedic surgeons, including the center of hip joint, the apex of femoral intercondylar fossa, the midpoint of tibial intercondylar ridge, the lowest point of medial and lateral femoral condyles, the nadir of medial and lateral tibial plateau and the midpoint of talus width, then the data set was established. High resolution network (HRNet) and transfer learning were used to construct the automatic detection model of key points, and the optimal model was selected with 5-fold cross validation. The coordinates of key points were detected, the key angles mechanical lateral distal femoral angle (mLDFA), medial proximal tibial angle (MPTA), joint line convergence angle (JLCA) and hip-knee-ankle angle (HKA) were calculated by cosine law, so as to realize the automatic measurement of key angles. The key angle automatic measurement model was constructed with automatic detection model of key points together with the key angles calculated by cosine law. On 50 randomly selected radiograms, the above mentioned key angles were measured by another 3 orthopedic doctors, then the consistency of the automatic detection model of key angles and the results measured by the physicians were evaluated. **Results** The mean values of mLDFA, MPTA, JLCA and HKA measured by 3 orthopedic surgeons was $(88.50 \pm 2.59)^\circ$, $(86.41 \pm 2.25)^\circ$, $(2.90 \pm 2.27)^\circ$ and $(174.62 \pm 3.97)^\circ$, respectively, those obtained with the automatic measurement model of key angles was $(88.48 \pm 2.60)^\circ$, $(86.52 \pm 2.57)^\circ$, $(3.11 \pm 2.41)^\circ$ and $(174.53 \pm 3.99)^\circ$, respectively, showing good agreement with the physician's measurement results ($ICC = 0.897, 0.888, 0.826, 0.996$). **Conclusion** The automatic measurement model of key angles on anteroposterior full-length lower limb radiograms based on DL was helpful for orthopedist to identify key points and to measure key angles.

[Keywords] genu varum; genu valgum; X-rays; lower limb alignment; key angles; deep learning; automatic measurement

DOI:10.13929/j.issn.1003-3289.2022.06.025

[基金项目] 国家重点研发计划(2018YFB1307802)。

[第一作者] 张子健(1996—),男,安徽宿州人,在读硕士。研究方向:医学图像处理。E-mail: 1621552359@qq.com

[通信作者] 马剑雄,天津大学天津医院骨科研究所,300050;天津市骨科生物力学与医学工程重点实验室,300050。E-mail: mjx969@163.com

[收稿日期] 2021-09-11 **[修回日期]** 2022-03-21

基于深度学习构建模型识别下肢全长正位 X 线片中的 下肢力线关键点及自动测量关键角度

张子健^{1,2}, 马剑雄^{1,2*}, 柏豪豪^{1,2}, 王 颖^{1,2}, 孙 磊^{1,2}, 卢 斌^{1,2},
高之浩^{1,2}, 于成双^{1,2}, 孙汉宸^{1,2}, 马信龙^{1,2}

(1. 天津大学天津医院骨科研究所, 天津 300050; 2. 天津市骨科生物力学与医学工程重点实验室, 天津 300050)

[摘要] **目的** 基于深度学习(DL)方法构建自动测量下肢全长正位 X 线片关键角度模型, 评估其临床应用价值。**方法** 回顾性选取 634 幅下肢全长正位 X 线片, 由 5 名骨科医师分别标注下肢力线关键点, 包括髌关节中心、股骨髁间窝顶点、胫骨髁间嵴中点、股骨内侧和外侧髁最低点、胫骨内侧和外侧平台最低点、距骨宽度中点, 并建立数据集。采用高分辨率网络(HRNet)进行迁移学习, 构建自动检测关键点模型, 以 5 折交叉验证筛选最优模型, 确定关键点坐标后, 通过余弦定律计算关键角度机械股骨远端外侧角(mLDFA)、胫骨近端内侧角(MPTA)、股骨胫骨关节线夹角(JLCA)及髌-膝-踝角(HKA), 实现自动测量关键角度, 并以关键点自动检测模型和通过余弦定律计算所得关键角度共同构建自动测量关键角度模型。随机选取 50 幅图像, 由另 3 名骨科医师手动测量上述关键角度, 评估自动测量关键角度模型与医师测量结果的一致性。**结果** 3 名骨科医师所测 mLDFA、MPTA、JLCA 及 HKA 的均值分别为 $(88.50 \pm 2.59)^\circ$ 、 $(86.41 \pm 2.25)^\circ$ 、 $(2.90 \pm 2.27)^\circ$ 及 $(174.62 \pm 3.97)^\circ$; 自动测量关键角度模型所获结果分别为 $(88.48 \pm 2.60)^\circ$ 、 $(86.52 \pm 2.57)^\circ$ 、 $(3.11 \pm 2.41)^\circ$ 及 $(174.53 \pm 3.99)^\circ$, 与医师测量结果的一致性较好 ($ICC=0.897, 0.888, 0.826, 0.996$)。**结论** 所构建的自动测量下肢全长正位 X 线片关键角度模型有助于识别骨科关键点和测量关键角度。

[关键词] 膝内翻; 膝外翻; X 线; 下肢力线; 关键角度; 深度学习; 自动测量

[中图分类号] R682.4; R445.4 **[文献标识码]** A **[文章编号]** 1003-3289(2022)06-0901-06

膝内、外翻畸形指双下肢自然伸直或站立时两膝(踝)相触而两踝(膝)不能同时闭合, 即下肢负重力线发生改变, 可致骨关节炎、髌骨软化等疾病^[1], 出现关节疼痛并影响关节功能。在下肢全长正位 X 线片上绘制下肢力线并测量关键角度是骨科治疗膝内、外翻畸形的关键^[2], 多需由骨科医师手动标点、连线并测量角度^[3], 存在主观误差且效率较低。深度学习(deep learning, DL)方法已广泛用于医学领域^[4-8], 基于 DL 自动测量下肢全长正位 X 线片中的关键角度的相关研究渐受关注, 如 PEI 等^[9]基于 DL 开发了自动测量下肢全长正位 X 线片中的髌-膝-踝角(hip-knee-ankel angle, HKA)系统; NGUYEN 等^[10]提出两阶段卷积神经网络(convolution neural network, CNN)算法, 用于测量下肢全长正位 X 线片中的关键角度。本研究基于 DL 方法构建自动测量下肢全长正位 X 线片关键角度模型, 并评估其临床价值。

1 资料与方法

1.1 一般资料 回顾性分析 2019 年 10 月—2021 年 2 月 634 例因膝关节疼痛于天津大学天津医院就诊患者, 男 254 例、女 380 例, 年龄 47~83 岁、平均 (66.4 ± 7.9) 岁; 膝内翻 255 例、外翻 221 例, 158 例无明显膝内翻或膝外翻。纳入标准: 负重位下肢全长正位 X 线

片资料完整, 同时清晰显示髌、膝及踝关节。排除关节置换术后、关节融合及图像质量差者。

1.2 仪器与方法 采用 Philips Digital Diagnost 平板式数字化 X 线摄影系统采集下肢全长正位 X 线片。嘱患者前后站立, 面向球管、背靠站立支架, 两足尖及内踝并拢, 尽量膝关节伸直; 曝光范围自髌关节至踝关节, 管电压及管电流机器自动调节, 焦片距 1.8 m, X 线球管自动旋转曝光 3 次, 以对接程序自动拼接图像。

1.3 关键点标注 分割图像后获得 1 268 幅左、右侧下肢全长正位图像。分别截取双侧髌、膝及踝关节图像, 由 5 名具有 10 年以上临床经验的骨科主治医师采用 Python 3.7 编写的程序分别标注下肢力线^[11]关键点, 包括髌关节中心、股骨髁间窝顶点、胫骨髁间嵴中点、股骨内侧和外侧髁最低点、胫骨内侧和外侧平台最低点、距骨宽度中点, 并依序设定为 O、A、B、C、D、E、F、G, 获得其坐标, 以标注坐标的均值为标签。见图 1。

1.4 建立模型 以各关节图像和标签构成数据集。基于网络高分辨率网络(high-resolution net, HRNet)^[12]以 DL 方法构建自动检测图像中的髌、膝及踝关节各关键点的模型。HRNet 定位关键点的精度主要取决于包含关键点信息的局部特征和整体图片



图1 标注下肢力线关键点及测量示意图

的全局语义特征及二者的信息融合。该网络由 4 个并行子网络构成,分为 4 个阶段,每个阶段引出下一层子网络,同一子网络的通道数不发生变化,并行子网络的通道数增加 1 倍,同时分辨率下降 50%,以使网络在始终保持高分辨率特征的同时获取充足的语义特征。两个阶段之间由融合模块连接,将不同分辨率的特征信息进行交换和融合,获得更好的关键点特征表示。

在 HRNet 基础上将 4 个子网络的低分辨率输出上采样至高分辨率输出,并连接起来作为高分辨率特征表示,获得 HRNetV2。根据髋、膝及踝关节需要检测的关键点数量,于 HRNetV2 后增加相应通道数的卷积层,作为输出层形成最终的卷积神经网络。见图 2。

基于 HRNet 进行迁移学习,训练时以 HRNetV2 在图像数据集分类任务中的预训练权重作为初始权重;经过高斯变换,以标注的关键点坐标标签为热图标签参与训练。训练好的模型可从各关节图像中检出包含关键点坐标信息的热图,表现为关键点坐标处颜色最

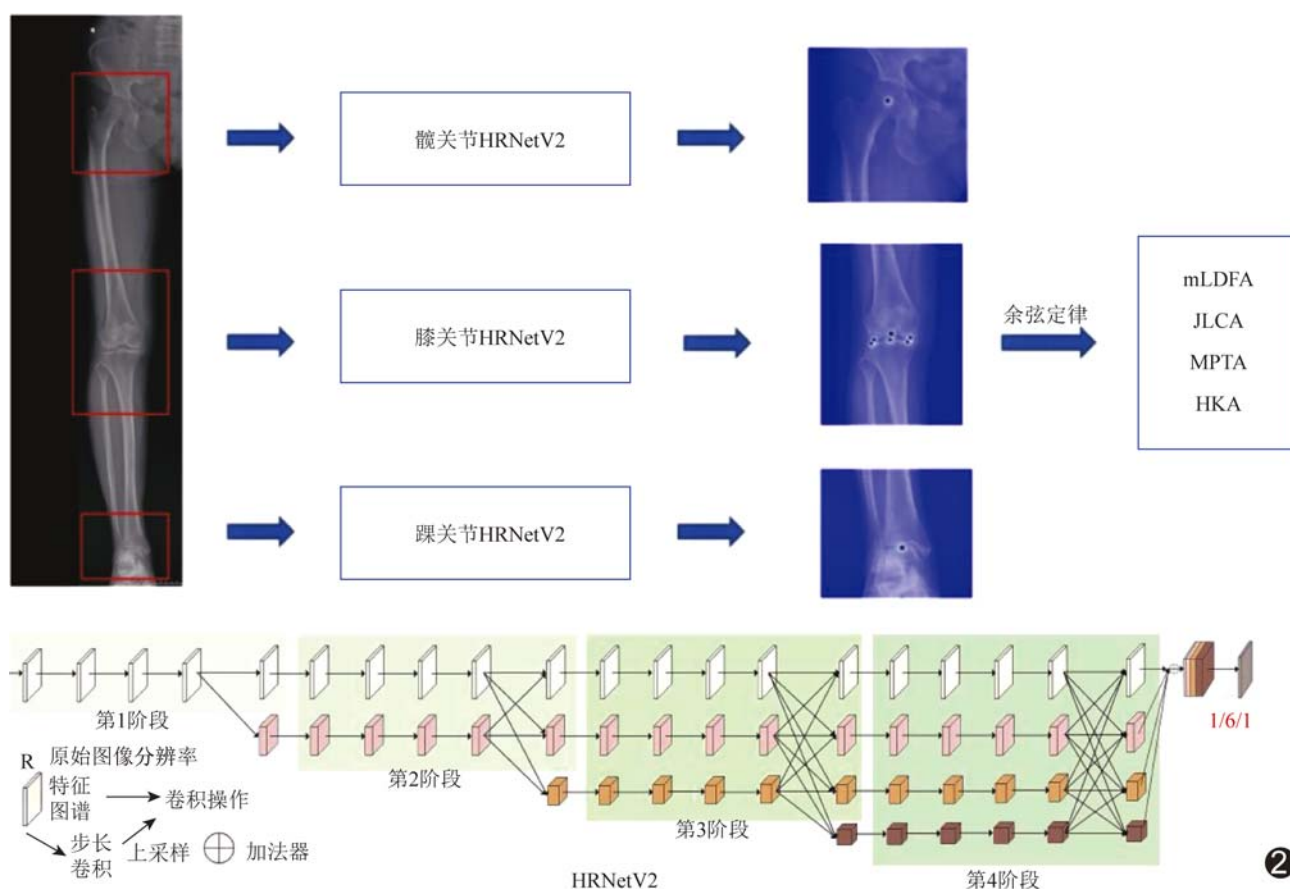


图2 基于 HRNet 模型自动检测髋、膝及踝关节关键点示意图

深、向外逐渐变浅;通过程序转换所获热图为关键点坐标值。采用 5 折交叉验证(图 3),将数据集随机均分无重复样本的 5 组,以其中一组为验证集、其余 4 组为训练集;重复以上步骤 5 次,轮流以各组样本作为验证集,计算自动检测模型在验证集上输出的关键点坐标与标签之间差值的平均值,以评估模型性能并筛选最优模型。

以自动检测模型检出髌、膝、踝关节最优关键点坐标后,通过余弦定律计算机械股骨远端外侧角(mechanical lateral distal femoral angle, mLDFA)(OA 与 CD 之间的夹角)、胫骨近端内侧角(medial proximal tibial angle, MPTA)(BG 与 EF 之间的夹角)、股骨胫骨关节线夹角(joint line convergence angle, JLCA)(CD 与 EF 之间的夹角)和 HKA(OA 与 BG 之间的夹角)(图 1);以自动检测关键点模型和经余弦定律计算所得关键角度共同构建自动测量关键角度模型。



图 3 5 折交叉验证示意图

1.5 模型效能验证及统计学分析 随机选取 50 幅下肢全长正位 X 线片,由另 3 名具有 30 年以上骨科经验的主任医师分别采用 AutoCAD 2019 软件(Autodesk 公司,美国)手动标注关键点,并测量 mLDFA、JLCA、MPTA、HKA,取其均值为最后结果,用于验证关键角度自动测量模型测量各角度的效能。

采用 SPSS 19.0 统计分析软件。以 $\bar{x} \pm s$ 表示计量资料。采用 Kendall 系数(W)评估 3 名医师测量各角度结果的一致性, $W > 0.8$ 为一致性极高。以组内相关系数(intra-class correlation coefficient, ICC)分析自动模型测量关键角度与 3 名医师测量结果的一致性, $ICC > 0.75$ 为一致性较好,并以 Bland-Altman 图直观评估其差异分布。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 筛选最优关键点自动检测模型 5 组关键点自动检测模型输出髌、膝及踝关节关键点坐标与标签之间差值的平均值分别为 1.01、0.94 及 0.97 pt(表 1),分别以第 3 折模型参数、第 1 折模型参数及第 3 折模型参数为最优模型参数。

表 1 5 组关键点自动检测模型输出髌、膝及踝关节关键点坐标与标签之间差值(pt)

关节部位	组 1	组 2	组 3	组 4	组 5
髌关节	1.01	1.04	0.98	1.03	0.98
膝关节	0.93	0.94	0.94	0.95	0.94
踝关节	0.97	0.97	0.95	0.99	0.96

2.2 关键角度自动测量模型测量各角度的效能 3 名骨科医师测量 mLDFA、MPTA、JLCA、HKA 的一致性均极高(W 均 > 0.95 , P 均 < 0.05),其均值分别为 $(88.50 \pm 2.59)^\circ$ 、 $(86.41 \pm 2.25)^\circ$ 、 $(2.90 \pm 2.27)^\circ$ 、 $(174.62 \pm 3.97)^\circ$ 。关键角度自动测量模型获得上述角度分别为 $(88.48 \pm 2.60)^\circ$ 、 $(86.52 \pm 2.57)^\circ$ 、 $(3.11 \pm 2.41)^\circ$ 、 $(174.53 \pm 3.99)^\circ$,与医师测量结果的一致性较好, ICC 分别为 0.897 [95% CI (0.851, 0.929)]、0.888 [95% CI (0.838, 0.923)]、0.826 [95% CI (0.752, 0.879)]、0.996 [95% CI (0.994, 0.997)]。

Bland-Altman 图(图 4)显示,自动测量关键角度模型与医师测量结果的差值分布于 95% 一致性界限内(差值均值 $\pm 1.96 \times$ 差值的标准差),二者一致性较好,其测量 mLDFA、MPTA、JLCA 和 HKA 的差值分别为 $(-0.013 5 \pm 1.178 9)^\circ$ 、 $(0.112 6 \pm 1.143 3)^\circ$ 、 $(0.2126 \pm 1.3826)^\circ$ 和 $(-0.092 2 \pm 0.359 7)^\circ$ 。

3 讨论

下肢力线一般指通过髌关节中心和踝关节中心的轴线^[11],其关键点主要集中在髌、膝及踝关节。下肢力线和关键角度可经由髌关节中心、股骨髁间窝顶点、胫骨髁间嵴中点、股骨内侧和外侧髁最低点、胫骨内侧和外侧平台最低点、距骨宽度中点 8 个关键点的位置加以确定。识别下肢全长正位 X 线片中上述关键点的位置准确与否影响关键角度测量结果。本研究构建自动识别关键点模型,以获得关键点在下肢全长正位 X 线片中的坐标,进而计算各关键角度;分别截取髌、膝及踝关节图像作为关键点自动检测模型训练的样本数据集以减少冗余信息、降低模型计算量,提高训练速度,增加模型精度,所获模型用于测量耗时不足 1 s。

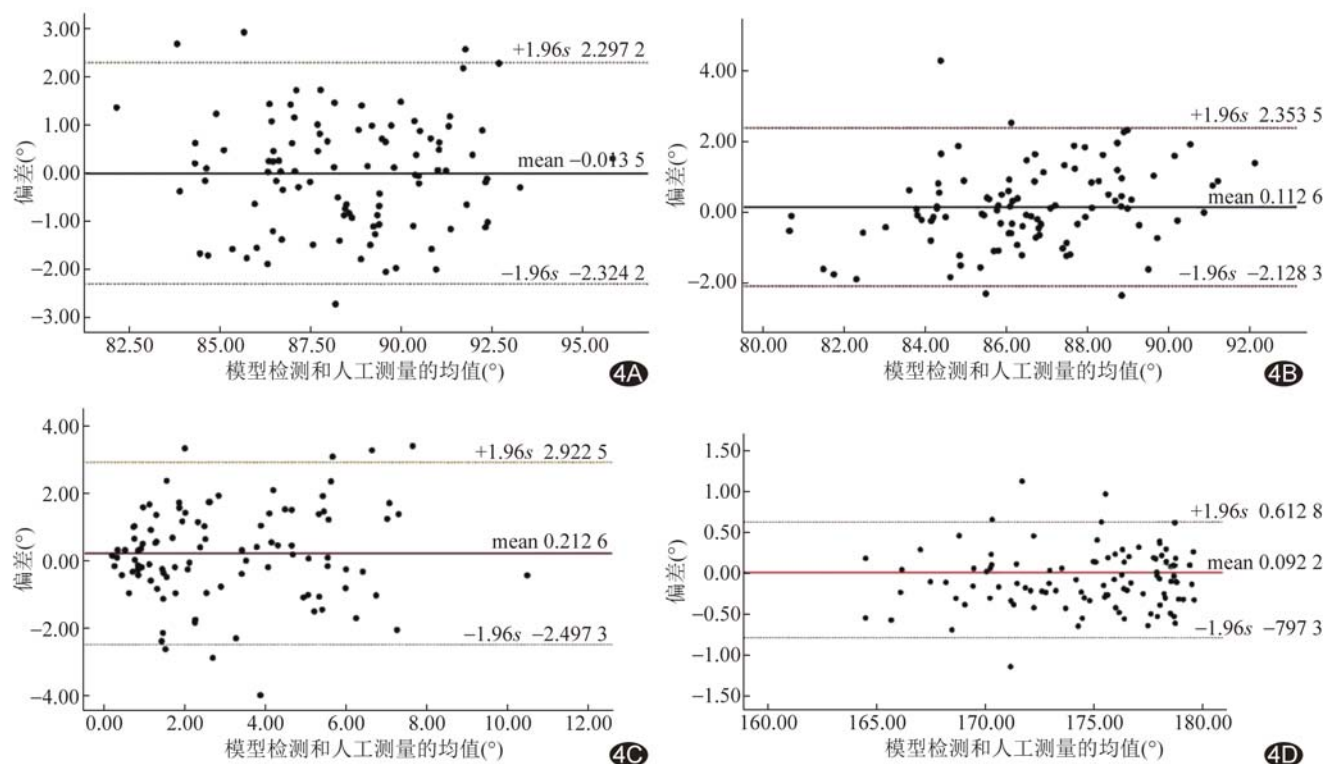


图 4 自动测量关键角度模型与医师测量结果的 Bland-Altman 图 A. mLDFA; B. MPTA; C. JLCA; D. HKA

在计算机视觉领域中,目前检测关键点相关研究主要集中于评估人体姿态(检测人体关键点)和人脸识别(检测人脸关键点)任务。为弥补空间精度损失,用于检测关键点的传统卷积神经网络常引入上采样或空洞卷积等,以提高特征分辨率,却并不能完全弥补空间分辨率的损失。本研究采用 HRNet 建立自动检测关键点模型,通过在并行多分辨率子网络上反复交换信息进行多尺度重复融合,始终维持高分辨率的表示,并通过迁移学习初始化模型权重,可在一定程度上弥补 DL 需要大量样本的缺陷,解决医学图像处理任务中数据集较小的问题^[13]。采用热图网络可直接回归出每一类关键点的概率,在一定程度上对每一个点均提供监督信息,使网络能够较快收敛;同时对每一个像素位置进行预测,能够提高定位关键点的精度。采用 5 折交叉验证,通过比较关键点自动检测模型输出的关键点坐标与标签之间的差值以筛选最优模型,提高模型对关键点检测精度。

本研究结果显示,自动测量关键角度模型与医师所测 mLDFA、MPTA、JLCA、HKA 的一致性均较好, HKA 的 ICC 高于 mLDFA 和 MPTA, JLCA 最小;且二者 HKA 测量值的差值小于 mLDFA 和 MPTA, JLCA 最大。分析原因,可能在于 JLCA 为关节线 CD

与 EF 的夹角, HKA 是股骨机械轴 OA 与胫骨机械轴 BG 的夹角, CD、EF 的长度远小于 OA 和 BG 的长度,故在检测关键点位置存在误差的情况下, JLCA 所受影响大于 HKA。

综上,以 DL 方法基于下肢全长正位 X 线片构建的自动测量关键角度模型有助于骨科医师识别关键点、测量关键角度,提高临床工作效率。本研究的主要局限性:①施行矫正手术时,需要同时考虑冠状位和矢状位下肢力线^[14],而本研究未分析矢状位所见下肢力线和关键角度;②为单中心研究,可能存在数据偏移;③通过人工标注获得训练标签,存在主观误差。

[参考文献]

- [1] MacDESSIS J, GRIFFITHS-JONES W, HARRIS I A, et al. Coronal plane alignment of the knee (CPAK) classification[J]. Bone Joint J, 2021, 103-B(2):329-337.
- [2] LIU X Y, CHEN Z, GAO Y, et al. High tibial osteotomy: Review of techniques and biomechanics[J]. J Health Eng, 2019, 2019:8363128.
- [3] CHO Y, SHIN H K, KIM E, et al. Postoperative radiologic outcome comparison between conventional and computer-assisted navigation total knee arthroplasty in extra-articular tibia vara[J]. J Orthop Surg (Hong Kong), 2020, 28(1):2309499020905702.

- [4] AHMAD O F, SOARES A S, MAZOMENOS E, et al. Artificial intelligence and computer-aided diagnosis in colonoscopy: Current evidence and future directions[J]. Lancet Gastroenterol Hepatol, 2019, 4(1):71-80.
- [5] WANI I M, ARORA S. Computer-aided diagnosis systems for osteoporosis detection: A comprehensive survey [J]. Med Biol Eng Comput, 2020, 58(9):1873-1917.
- [6] AL-ANTARI M A, AL-MASNI M A, KIM T S. Deep learning computer-aided diagnosis for breast lesion in digital mammogram[J]. Adv Exp Med Biol, 2020, 1213:59-72.
- [7] NAYANTARA P V, KAMATH S, MANJUNATH K N, et al. Computer-aided diagnosis of liver lesions using CT images: A systematic review[J]. Comput Biol Med, 2020, 127:104035.
- [8] AHMED S, HOSSAIN T, HOQUE O B, et al. Automated COVID-19 detection from chest X-ray images: A high-resolution network (HRNet) approach [J]. SN Comput Sci, 2021, 2(4):294.
- [9] PEI Y, YANG W, WEI S, et al. Automated measurement of hip-knee-ankle angle on the unilateral lower limb X-rays using deep learning[J]. Phys Eng Sci Med, 2021, 44(1):53-62.
- [10] NGUYEN T P, CHAE D S, PARK S J, et al. Intelligent analysis of coronal alignment in lower limbs based on radiographic image with convolutional neural network [J]. Comput Biol Med, 2020, 120:103732.
- [11] CAO Z, HIDALGO G, SIMON T, et al. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields [J]. IEE Trans Pattern Anal Mach Intell, 2021, 43(1):172-186.
- [12] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation[J/OL]. Los Alamos: arXiv e-prints, 2019 (2019-02-25) [2021-08-25]. <https://arxiv.org/pdf/1902.09212.pdf>.
- [13] 唐思源, 刘燕茹, 杨敏. 基于迁移学习和三维卷积神经网络检测肺结节[J]. 中国医学影像技术, 2020, 36(12):1882-1886.
- [14] SCHOENMAKERS D A L, FECZKO P Z, BOONEN B, et al. Measurement of lower limb alignment: There are within-person differences between weight-bearing and non-weight-bearing measurement modalities [J]. Knee Surg Sports Traumatol Arthrosc, 2017, 25(11):3569-3575.

使用阿拉伯数字和汉字数字的一般原则

根据 GB/T 15835《出版物上数字用法的规定》

(1) 在统计图表、数学运算、公式推导中所有数字包括正负整数、小数、分数、百分数和比例等,都必须使用阿拉伯数字。

(2) 在汉字中已经定型的词、词组、成语、缩略语等都必须使用汉语数字,例如:一次方程、三维超声、二尖瓣、法洛四联症、星期一、五六天、八九个月、四十七八岁等。

(3) 除了上述情况以外,凡是使用阿拉伯数字而且又很得体的地方,都应该使用阿拉伯数字。遇到特殊情况时,可以灵活掌握,但应该注意使全篇同一。

(4) 如果数字的量级小于 1 时,小数点前面的零(0)不能省去,如 0.32 不能写成.32。