

基于匹配线索回归的侧面脊柱关键点检测

高孟豪¹, 郭立君^{1*}, 张荣¹, 倪丽欣^{2,3}, 王强⁴, 何秀超⁴¹宁波大学信息科学与工程学院, 浙江 宁波 315211;²宁波大学医学院, 浙江 宁波 315211;³宁波市海曙区第二医院, 浙江 宁波 315099;⁴宁波大学附属第一医院, 浙江 宁波 315000

摘要 在检测侧面脊柱关键点时, 由于受到器官遮挡的影响, 以往的热图回归方法难以区分不同椎骨上的关键点, 容易出现关键点与对应椎骨的匹配错误。为了解决这个问题, 提出了一个新的单阶段侧面脊柱关键点检测方法, 该方法同时预测关键点热图和关键点匹配线索(椎骨中心热图和关键点 offset), 利用匹配线索建立关键点与对应椎骨的匹配关系。为了提升匹配效果, 提出几何感知特征增强模块, 通过提取关键点特征增强椎骨中心的特征表达。此外, 利用加权损失函数缓解关键点热图和椎骨中心热图中正负样本比例失衡问题。实验结果表明, 所提方法的平均检测误差为 8.84, 相较于性能第二的方法精度提升 36%。

关键词 医用光学; 关键点检测; 卷积网络; 可变形卷积; 脊柱侧弯

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP231172

Lateral Spine Landmark Detection Based on Matching Clue Regression

Gao Menghao¹, Guo Lijun^{1*}, Zhang Rong¹, Ni Lixin^{2,3}, Wang Qiang⁴, He Xiuchao⁴¹Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, Zhejiang, China;²School of Medicine, Ningbo University, Ningbo 315211, Zhejiang, China;³Haishu District Second Hospital of Ningbo, Ningbo 315099, Zhejiang, China;⁴The First Affiliated Hospital of Ningbo University, Ningbo 315000, Zhejiang, China

Abstract In lateral spine landmark detection, the previous heatmap regression methods have difficulty in distinguishing landmarks on different vertebrae due to the influence of organ occlusion and are prone to landmark and vertebrae matching errors. To solve this problem, we propose a new one-stage lateral spine landmark detection method, which simultaneously predicts the landmark heatmap and landmark matching clue (vertebra center heatmap and landmark offset), and uses the matching clue to match the landmarks with the corresponding vertebra. In order to improve the matching effect, we propose the geometry-aware feature aggregator module, which can extract the landmark features on the vertebra to enhance the feature representation of the vertebra center. We also use a weighted loss function to alleviate the imbalance of positive and negative samples in the landmark and the vertebra center heatmaps. Experimental results show that the average detection error of the proposed method is 8.84, which has 36% improvement in accuracy compared to the method with the second-highest performance.

Key words medical optics; landmark detection; convolutional neural networks; deformable convolution; scoliosis

1 引言

特发性脊柱侧弯是一种严重的疾病, 严重影响患者的心肺健康^[1]。临床上, 医生通常利用正侧面脊柱 X 线片上的关键点估计 Cobb 角、胸椎后凸角(TK)和

腰椎前凸角(LL)等参数来量化评估脊柱侧弯^[1-3]。精确的正侧面脊柱关键点检测对于量化评估脊柱侧弯具有重要的意义。由于侧面脊柱图像数据量少且存在严重的器官及肋骨遮挡现象, 精确的侧面脊柱关键点检测更具挑战性。

收稿日期: 2023-04-25; 修回日期: 2023-05-12; 录用日期: 2023-05-29; 网络首发日期: 2023-05-31

基金项目: 浙江省自然科学基金/公益技术项目(LGF21F020008)、宁波市公益性科技计划项目(2022S134)

通信作者: *guolijun@nbu.edu.cn

目前,脊柱关键点检测的方法主要分为两类:两阶段法和单阶段法。两阶段法首先通过椎骨检测确定每个椎骨的范围,然后在椎骨范围内对单个椎骨进行关键点检测。两阶段方法^[4-6]鲁棒性较好但存在不能端到端训练和无法利用图像全局信息等问题。单阶段法通常直接对脊柱图像中的关键点进行检测,具有可端到端训练和可利用图像全局信息等优点。单阶段法主要分为三类:坐标回归方法、中心加偏移方法和热图回归方法。坐标回归方法^[7-8]通常利用网络直接预测关键点坐标,该方法受限于参数数量和输入图像尺寸,精度不高^[9];中心加偏移方法^[9-10]通过预测椎骨中心和关键点与对应椎骨中心的偏移(offset)获得关键点坐标,该方法容易受到椎骨中心预测精度的限制,进而导致精度不足;热图回归方法^[11-13]通过回归关键点热图并使用非极大值方法抑制获得关键点坐标,该方法相较于坐标回归方法具有更好的鲁棒性和精度。然而由于不同椎骨上相同位置的关键点具有相似的特征,热图回归方法难以区分不同椎骨上的关键点,容易造成椎骨与关键点的匹配错误,如,第一节椎骨上的关键点被误识别为第二节椎骨上的关键点。在侧面脊柱关键点检测中,侧面脊柱 X 线片存在严重的器官和肋骨遮挡,导致匹配错误更加严重。匹配错误对于后续任务如脊柱

三维重建、TK 和 LL 评估等有一定的影响。

为了解决上述问题,本文提出了一个新的单阶段侧面脊柱关键点检测方法,该方法同时预测关键点热图和关键点匹配线索(椎骨中心热图和关键点 offset),利用关键点匹配线索建立关键点与对应椎骨的匹配关系。当椎骨的某一部分被遮挡时,该椎骨上未被遮挡的关键点特征有助于椎骨中心的预测。为了缓解遮挡对于椎骨中心预测的影响,提出几何感知特征增强(GFA)模块,该模块通过采样关键点特征增强椎骨中心的特征表达。为了缓解关键点热图和椎骨中心热图中正负样本比例失衡问题,提高关键点热图和椎骨中心热图的预测质量,对关键点热图和椎骨中心热图使用加权损失函数进行监督训练。

2 方法介绍

所提侧面脊柱关键点检测网络如图 1 所示,包含骨干(backbone)网络和两个检测分支:关键点分支和匹配线索分支。将 ResNet34 和 FPN(feature pyramid networks)作为 backbone 网络提取图像特征获得特征图(feature map),在得到 feature map 之后分别送入关键点分支和匹配线索分支,其中: H 、 W 分别为输入图像的高和宽; K 为关键点类别数。

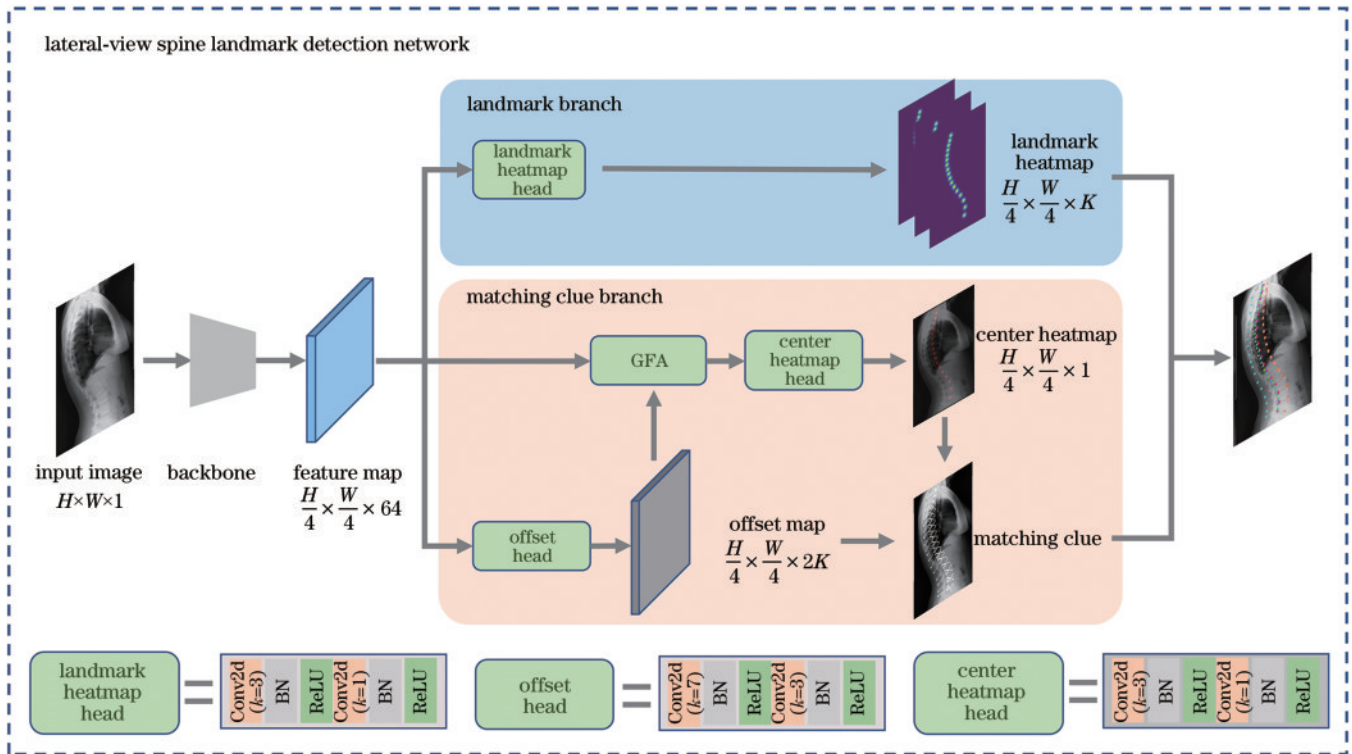


图 1 所提侧面脊柱关键点检测方法框架

Fig. 1 Proposed lateral spine landmark detection method framework

在关键点分支中,feature map 经过 landmark heatmap head 获得预测的关键点热图。landmark heatmap head 包含卷积层、BN (batch normalization) 层和 ReLU (rectified layer unit)层,其中两个卷积层的输入通道分

别为 64 和 256,卷积核大小 k 分别为 3 和 1,步长均为 1。在匹配线索分支中,feature map 首先经过 offset head 得到预测的关键点 offset map。然后 feature map 和 offset map 被送入 GFA 模块增强椎骨中心的特征表

达,之后经过 center heatmap head 得到预测的椎骨中心热图。其中 center heatmap head 与 landmark heatmap head 参数设置相同,offset head 与 landmark heatmap head 结构类似但两个卷积层的卷积核更大,分别为 7 和 3,这样可以更好地学习关键点与椎骨中心的关系。在推理阶段,首先从预测的椎骨中心热图与 offset map 中获得关键点匹配线索,其次从关键点热图中提取关键点坐标,最后使用匹配方法建立关键点与椎骨的匹配关系。

2.1 关键点热图

考虑到椎骨之间的相似性,即每节椎骨相同位置的关键点具有类似的特征,按照脊柱关键点的位置和临床定义将脊柱关键点分为 9 类,同一类关键点用同一热图进行监督。坐标为 (x, y) 的关键点在输出热图中的坐标为 $(x/n, y/n)$,其中, n (此处为 4) 为输出热图相较于输入图像的下采样倍数。该关键点的真实值 (ground-truth) 使用一个高斯核 G 表示,热图中 (i, j) 处的值的计算方式为

$$G(i, j) = \exp\left[-\frac{(i - \lfloor x/n \rfloor)^2 + (j - \lfloor y/n \rfloor)^2}{2\sigma^2}\right], \quad (1)$$

式中: σ (默认为 5) 为高斯核半径; $\lfloor \cdot \rfloor$ 表示向下取整。对属于同一类的关键点热图,通过在相同位置取最大值的方式合并成一个单通道热图。

2.2 椎骨中心热图

与关键点热图类似,使用高斯核 G_c 表示椎骨中心 ground-truth:

$$G_c(i, j) = \exp\left[-\frac{(i - \lfloor x_c/n \rfloor)^2 + (j - \lfloor y_c/n \rfloor)^2}{2\sigma_c^2}\right], \quad (2)$$

式中: (x_c, y_c) 为椎骨中心坐标,由椎骨 4 个角点 (椎骨上端板左缘、椎骨上端板右缘、椎骨下端板左缘、椎骨下端板右缘) 坐标取平均得到; σ_c 为 5; n 为 4。对于所有的椎骨中心热图,同样通过取最大值的方式合并为一个单通道热图。

2.3 关键点 offset

关键点 offset 是指从椎骨中心到该椎骨上各关键点的向量,表示为

$$\left(\frac{x - x_c}{n}, \frac{y - y_c}{n}\right), \quad (3)$$

式中: (x, y) 为关键点坐标; (x_c, y_c) 为对应的椎骨中心坐标; n 为 4。通过预测 offset map 获得关键点 offset。其中 offset map 通道数量为 $2 \times K$, K (本文中为 9) 为关键点类别数。

2.4 GFA 模块

椎骨上分布有若干个关键点,这些关键点在椎骨上呈现一定的几何结构,利用这些关键点的特征可以增强椎骨中心的特征表达,缓解遮挡对于椎骨中心预测的影响。因此提出 GFA 模块,该模块结合可变形卷积^[14-15]和 offset head 预测的 offset map,动态地提取关键点特征增强椎骨中心的表达。与以往的可变形卷积采用间接监督的方式预测 offset map 不同, GFA 模块采用直接监督的方式预测 offset map,可以更加高效地指导网络学习关键点的几何结构。

GFA 模块的示意图如图 2 所示,该模块的输入为 feature map 和关键点 offset map。其中, offset map 中包含网络预测的关键点与对应椎骨中心的 offset。对于输入 feature map,根据 offset map 中对应位置的 offset 采样 9 个位置的信息得到新的 feature map,然后将采样后的 feature map 与输入 feature map 拼接 (Concat) 并输出。

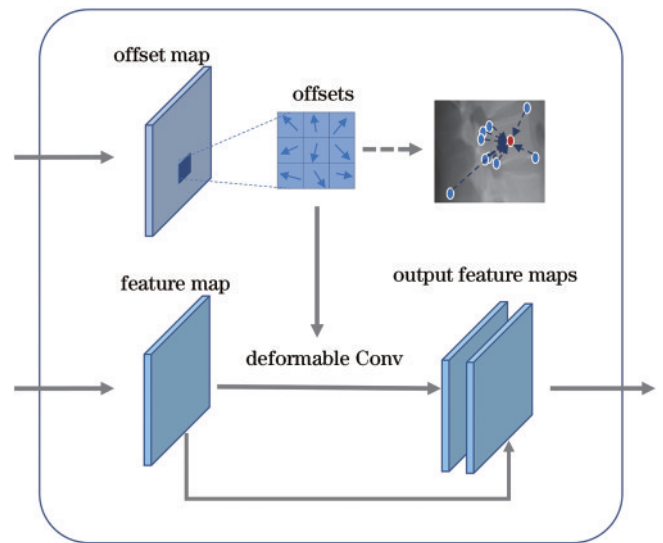


图 2 GFA 模块
Fig. 2 GFA module

具体地,对于输入 feature map 的每个位置,通过可变形卷积 (3×3 卷积) 和 9 个关键点与对应椎骨中心的 offset 进行采样得到新的特征:

$$\hat{F}(p_i) = \sum_{m=1}^9 w_m \cdot F(p_i + \Delta p_i^m), \quad (4)$$

式中: $\hat{F}(p_i)$ 为输入 feature map 中第 i 个位置采样后的特征; w_m 为卷积核的权重; p_i 为 feature map 中第 i 个位置的坐标; Δp_i^m 为第 m 个关键点与第 i 个位置的 offset; $F(p_i)$ 表示 p_i 处的图像特征。

2.5 关键点匹配方法

为了获得脊柱关键点的精确坐标,对于输入的一张侧面脊柱图片,同时预测关键点热图和关键点匹配线索。对于关键点热图,利用非极大值抑制 (NMS) 方法获得若干候选关键点构成的集合, $S = \{S_1, S_2, \dots, S_K\}$ 。其中, S_j 为第 j 类的候选关键点集合,

该集合包含若干(本文为 25 个)第 j 类的候选关键点。

对于椎骨中心热图和关键点 offset, 首先使用 NMS 获得 17 个预测的椎骨中心坐标, 然后将每个预测的椎骨中心坐标与关键点 offset 相加获得每节椎骨上数量为 K 的关键点的匹配线索 $g = \{g_1, g_2, \dots, g_{17}\}$, 其中 g_i 包含第 i 节椎骨的数量为 K 的关键点的初始位置坐标。

在获得候选关键点集合 S 与关键点匹配线索 g 之后, 利用关键点匹配线索建立关键点与对应椎骨的匹配关系。具体地, 对于第 i 节椎骨的第 j 个关键点 p_{ij} , 其初始位置坐标为 g_{ij} , 选择第 j 类候选关键点集合 S_j 中与 g_{ij} 距离最近的点作为点 p_{ij} 的预测值。

2.6 损失函数

所提损失函数分为两部分: offset 损失函数和热图损失函数。使用 L_1 损失函数监督 offset; 为了缓解关键点热图和椎骨中心热图中正负样本不平衡问题, 使用加权损失函数监督训练关键点热图和椎骨中心热图, 加权损失函数表示为

$$L_h = \|\mathbf{M} \odot (H - H^*)\|_2^2, \quad (5)$$

式中: \mathbf{M} 为权重矩阵, 尺寸大小与热图相同, 该权重矩阵中关键点范围内的值为 1, 其余位置均为 0.1; \odot 表示点乘; H 为预测热图; H^* 为 ground-truth; $\|\cdot\|_2^2$ 为二范数。

总损失为 offset 损失与热图损失之和:

$$L_{\text{total}} = L_o + L_h^c + L_h^l, \quad (6)$$

式中, L_o 、 L_h^c 、 L_h^l 分别为 offset 损失、椎骨中心热图损失和关键点热图损失。

3 数据集及评价指标

3.1 数据集

构建的新的侧面脊柱关键点检测数据集包含 328 例侧面脊柱 X 线片图像, 该数据集的图像及标注如图 3 所示, 每张图像包含从胸椎第 1 节(T1)到腰椎第 5 节(L5)共计 17 节椎骨。与以往的数据集只标注

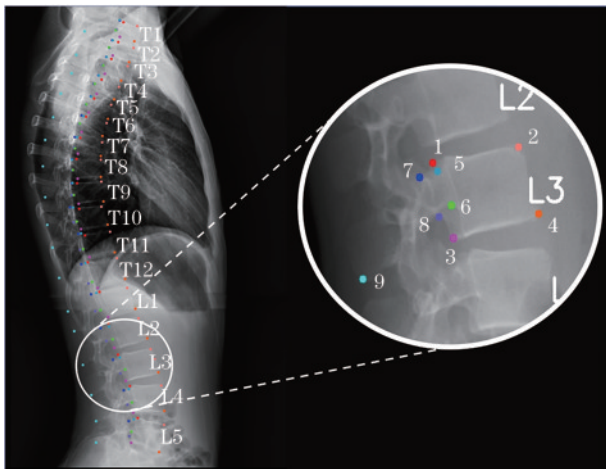


图 3 侧面脊柱关键点数据集

Fig. 3 Lateral spine X-rays dataset

每节椎骨 4 个关键点不同, 构建的数据集中每节椎骨上标注了椎骨上端板左缘、椎骨上端板右缘、椎骨下端板左缘、椎骨下端板右缘、右侧椎弓根上缘、右侧椎弓根下缘、左侧椎弓根上缘、左侧椎弓根下缘和棘突共 9 个关键点(图 3 放大图中标号 1~9 代表不同类的脊柱关键点)。每张图像共包含 153 个关键点。以上标注均由多名资深影像科医生和脊柱外科医生完成并进行交叉验证以确保可靠性。每张 X 线片的图像像素距离(image pixel spacing)为(0.32, 0.32), 单位为 mm。图像尺寸约为 2198 pixel×1128 pixel。该数据集不仅可以用于侧面脊柱参数估计(如 TK 和 LL), 还可用于其他参数估计任务, 因此该数据集具有一定的临床意义和价值。

实验过程中, 使用 60% 的图像作为训练集, 20% 的图像作为验证集, 20% 的图像作为测试集。

3.2 实验设置

实验均使用 Ubuntu20.04 平台和一张 NVIDIA 3090 GPU 完成。Backbone 网络为 ResNet34 和 FPN, ResNet34 使用 ImageNet^[16]数据集预训练, 其他网络权重均使用标准高斯分布初始化。训练时输入图像均缩放到 1024 pixel×512 pixel, 输出热图尺寸为 512 pixel×128 pixel。为了避免过拟合, 采用了数据增强, 包括裁剪和对对比度扭曲。网络优化器为 Adam, 初始学习率为 2.5×10^{-4} 。每次实验训练 50 个 epoch, 当验证集上的损失没有明显下降时停止训练。

3.3 评价指标

为了评估脊柱关键点检测的精度, 将平均检测误差 E_{dec} ^[9]、归一化的平均绝对误差(Scaled MAE)^[17]和均方误差(MSE)^[7]作为评价指标, 具体计算公式为

$$E_{\text{dec}} = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - x_i^*)^2 + (y_i - y_i^*)^2}, \quad (7)$$

$$E_{\text{scaled MAE}} = \frac{1}{2N} \sum_{i=1}^N \left(\frac{|x_i - x_i^*|}{w} + \frac{|y_i - y_i^*|}{h} \right), \quad (8)$$

$$E_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N \left[\left(\frac{x_i - x_i^*}{w} \right)^2 + \left(\frac{y_i - y_i^*}{h} \right)^2 \right], \quad (9)$$

式中: (x_i, y_i) 为预测值; (x_i^*, y_i^*) 为 ground-truth; N 为脊柱关键点总个数; w, h 为测试图像的宽和高。计算评价指标时, 预测值与 ground-truth 均还原到原始图像坐标系, 单位为 pixel。

为了评估 TK 和 LL 的精度, 将对称平均绝对误差(SMAPE)^[18]作为评价指标, 表示为

$$E_{\text{SMAPE}} = \sum_{i=1}^M \left(\frac{|a_i - b_i|}{|a_i + b_i|} \right), \quad (10)$$

式中: a 和 b 分别为利用脊柱关键点(上下端板的端点)的预测值和 ground-truth 计算得到的角度; M 为测试图像个数。

4 分析与讨论

4.1 对比实验

为了验证所提方法的性能,与目前先进的两阶段脊柱关键点检测方法进行了对比实验。由于棘突与邻近的椎骨较为靠近,椎骨边界框很难将其包含在内,为了公平起见,仅比较每节椎骨上除棘突外其余 8 个关键点的检测精度。

从表 1 可以看出,两阶段方法在 8 个关键点的检测中表现较好,这主要是由于两阶段方法中第一阶段获得的椎骨边界框为关键点检测提供了位置约束,避免了背景噪声对于关键点检测的影响。此外,第二阶段只对单个椎骨进行关键点检测,避免了匹配错误。所提方法相较于两阶段方法在精度上表现更好,这主要是由于所提方法可以利用图像的全局信息,而两阶段方法将椎骨区域进行切分然后独立地进行关键点检测,无法利用全局信息。

对于单阶段方法,对比了包括棘突在内的 9 个关键点的检测效果,实验结果如表 2 所示。所提方法在精度上优于其他单阶段方法。所提方法与其他单阶段方法的对比如图 4 所示。其中:三角形为预测值;圆点为真实值;直线连接真实值与预测值,直线越长表示偏

表 1 所提方法与两阶段方法对比实验结果

Table 1 Comparative experiment results between two-stage methods and the proposed method

Method	E_{dec}	Scaled MAE	MSE
Khanal et al. [4] proposed method	17.76	0.007407	0.000235
Chen et al. [6] proposed method	23.36	0.007580	0.000472
Zhang et al. [5] proposed method	11.01	0.004187	0.000122
Ours	7.14	0.003386	0.000060

表 2 所提方法与单阶段方法对比实验结果

Table 2 Comparative experiment results between one-stage methods and the proposed method

Method	E_{dec}	Scaled MAE	MSE
Yi et al. [9] proposed method	13.89	0.005965	0.000321
Zhang et al. [13] proposed method	23.83	0.009693	0.001639
Ao et al. [11] proposed method	48.84	0.018596	0.025107
Ours	8.84	0.004147	0.000111

离真实值越远。在图像较为清晰时(case 1),脊柱关键点均能被正确地检测出来且误差较小(直线较短)。在图像存在轻微遮挡时(case 2):Zhang 等[13]和 Ao 等[11]的方法对于边缘不清晰的椎骨均出现了匹配错误(直线

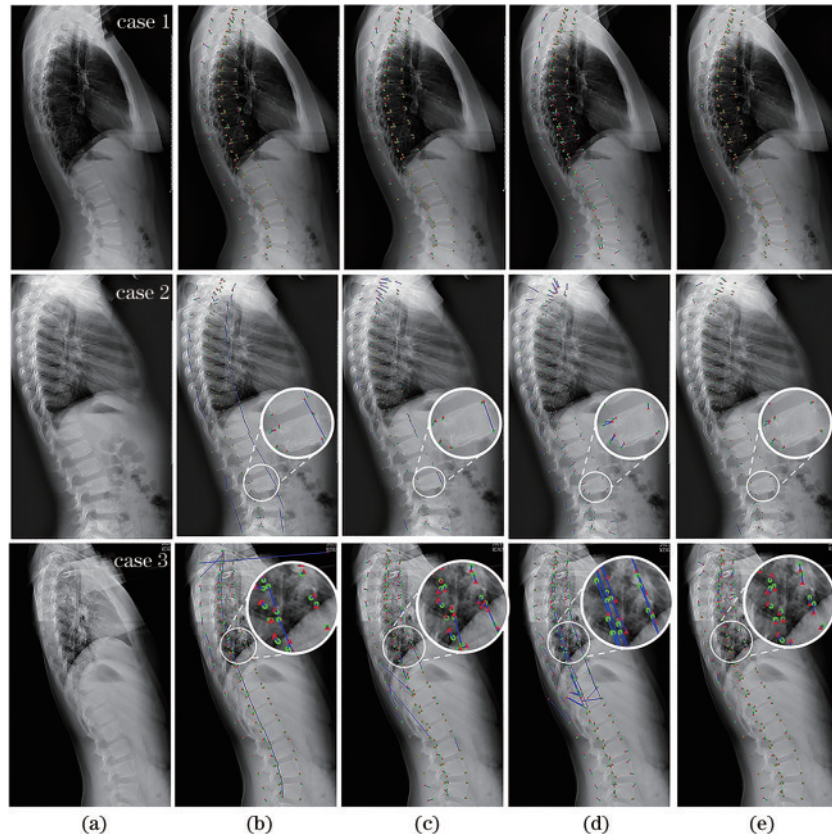


图 4 与其他单阶段脊柱关键点检测方法的对比结果。(a)输入图像;(b)Zhang 等[13]方法检测结果;(c)Ao 等[11]方法检测结果;(d)Yi 等[9]方法检测结果;(e)所提方法检测结果

Fig. 4 Comparison results between other single-stage methods and proposed method. (a) Input image; (b) results by Zhang et al. [13] proposed method; (c) results by Ao et al. [11] proposed method; (d) results by Yi et al. [9] proposed method; (e) results by our method

较长且连接到了另一关键点);Yi等^[9]的方法通过椎骨中心和 offset 定位关键点,可以有效减少匹配错误;相较于 Yi等^[9]的方法,所提方法的预测结果精度更高(直线更短)。在图像存在较为严重的遮挡时(case 3):Zhang等^[13]和Ao等^[11]方法的预测结果均存在部分距真实值较远(直线较长)的异常值和较多匹配错误;Yi等^[9]方法由于某个椎骨中心的预测错误,部分椎骨的关键点全部匹配错误。相较于上述方法,所提方法一方面通过 GFA 模块和加权损失函数提高椎骨中心的预测质量,提高关键点匹配效果;另一方面利用关键点热图获得更加精确的关键点坐标,因此实现了更好的检测效果。

为了验证所提方法在侧面参数评估中的效果,还比较了所提方法与其他先进方法在 TK 和 LL 估计中的精度。评价指标为 SMAPE,值越小表示性能越好。其中 TK 和 LL 均使用预测的脊柱关键点(上下端板端点)计算得到。实验结果如表 3 所示。由表 3 可知,所提方法相较于其他方法在 TK 和 LL 估计中具有较好

表 3 所提方法与其他先进方法在 TK 和 LL 估计中的 SMAPE 对比结果

Table 3 Comparison results of SMAPE between the proposed method with other advanced methods in TK and LL estimation

Method type	Method	TK /%	LL /%
Two-stage	Khanal et al. ^[4] proposed method	37.22	36.32
	Chen et al. ^[6] proposed method	22.79	19.45
	Zhang et al. ^[5] proposed method	12.83	10.53
One-stage	Yi et al. ^[9] proposed method	8.42	5.91
	Zhang et al. ^[13] proposed method	8.63	5.55
	Ao et al. ^[11] proposed method	17.53	9.49
	Ours	8.40	3.62

的效果。相较于两阶段法,由于所提方法可以利用全局信息,关键点检测精度更高,因此 TK 和 LL 估计精度更高;相较于单阶段方法,由于所提方法减少了匹配错误且精度更高,因此在 TK 和 LL 估计中表现较好。

4.2 消融实验

为了验证所提方法的有效性,进行了消融实验,评价指标为 E_{dec} 。实验结果如表 4 所示。其中:“baseline”表示去除了加权损失函数和 GFA 模块;“center”表示椎骨中心;“landmark”表示使用热图回归与匹配线索;“center+offset”表示使用椎骨中心加关键点 offset;“weighted loss”表示使用加权损失函数。

表 4 所提方法的消融实验结果
Table 4 Ablation experiment results of the proposed method

Baseline	Weighted loss	GFA	Center	Center +offset	Landmark
✓			9.63	14.03	11.75
✓	✓		8.28	12.65	10.86
✓		✓	6.45	10.73	9.60
✓	✓	✓	6.21	10.12	8.84

从表 4 可以看出,相较于 baseline,所提方法在检测精度上均有一定的提升。其中:使用加权损失函数可以提高关键点和椎骨中心的预测精度;所提 GFA 模块能够有效提升椎骨中心的预测精度,提高关键点匹配效果;如图 5 所示,使用加权损失函数后,腰椎部分的直线更短,这说明匹配错误更少;同时使用加权损失函数和 GFA 模块后匹配错误进一步减少。此外,从表 4 可知,中心加偏移的方法在预测精度上低于热图回归方法,这是由于热图回归方法可以直接定位关键点,而中心加偏移的方法间接定位关键点,会同时受到椎骨中心预测误差和偏移预测误差的影响。

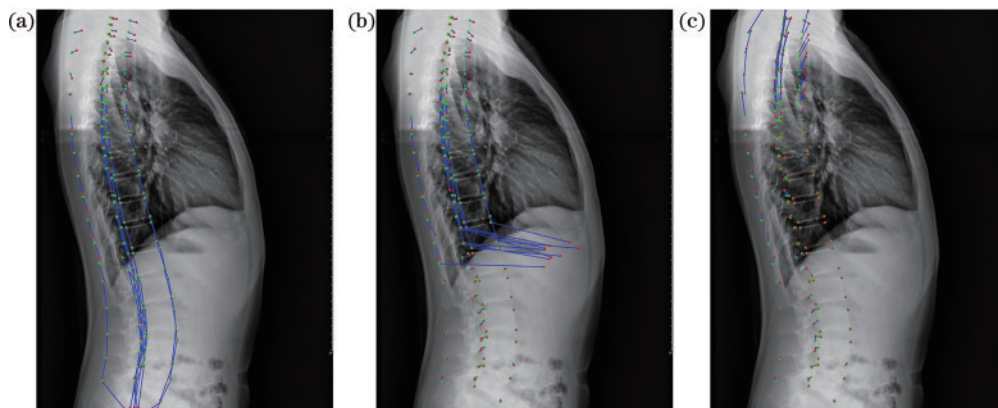


图 5 消融实验结果。(a)基线;(b)使用加权损失函数;(c)使用加权损失函数和 GFA 模块;
Fig. 5 Ablation experiment results. (a) Baseline; (b) with weighted loss; (c) with weighted loss and GFA module

为了比较所提 GFA 模块与以往的可变形卷积的性能,进行了一组对比实验。实验结果如表 5 所示。其中:“DCN”表示采用可变形卷积的间接监督方式训练采样 offset;“GFA”表示采用直接监督方式训练采

样 offset。从实验结果可知,采用直接监督的方式训练采样 offset 可以更加高效地学习关键点与椎骨中心的几何结构,以及利用关键点特征增强椎骨中心的预测质量,而可变形卷积采用间接监督的方式效果不佳。

表 5 所提 GFA 模块与可变形卷积的对比结果

Table 5 Comparison results between proposed GFA module with deformable convolution

Method	Center	Center+offset	Landmark
DCN	8.56	12.12	10.68
GFA	6.45	10.73	9.60

为了验证匹配线索的作用,对比了使用匹配线索和不使用匹配线索的检测效果。实验结果如表 6 所示。其中:“w/ matching clue”表示使用匹配线索;“w/o matching clue”表示不使用匹配线索。不使用匹配线索时,对每类脊柱关键点热图使用 NMS 方法可获得 17 个关键点预测坐标,然后按照纵坐标排序并按照次序与椎骨匹配。从实验结果可知,使用匹配线索时误差更小,这主要是由于匹配线索有助于匹配关键点与对应椎骨,减少匹配错误。此外,匹配线索还对关键点起到一定的约束作用,可以去除异常值。

表 6 所提关键点匹配方法的有效性

Table 6 Validity of landmark matching method

Method	Landmark
w/o matching clue	12.20
w/ matching clue	8.84

5 结 论

在侧面脊柱关键点检测中,针对以往热图回归方法中存在的匹配错误问题,提出了基于匹配线索回归的方法。该方法在预测关键点热图的同时,通过预测椎骨中心热图和关键点 offset 为关键点提供匹配线索,从而大大降低了匹配错误。为了缓解遮挡对于椎骨中心预测的影响,获得更好的匹配效果,使用 GFA 模块采样关键点特征增强椎骨中心的特征表达;利用加权损失函数缓解关键点热图和椎骨中心热图中正负样本不平衡问题。实验结果显示利用关键点特征可以提高椎骨中心热图的预测质量,采用直接监督的方式可以更加高效地学习关键点的几何结构。最终所提方法在侧面脊柱关键点检测中取得了较好的检测精度和鲁棒性,并在侧面临床参数估计中取得良好表现。然而,所提方法还存在一定的不足,如,当某节椎骨被完全遮挡或连续几节椎骨都被遮挡时,椎骨中心仍然会出现较大偏差导致关键点匹配错误。因此,解决大面积遮挡问题将是未来工作的重点。

参 考 文 献

[1] Zhang K L, Xu N F, Guo C Y, et al. MPF-net: an effective framework for automated Cobb angle estimation [J]. *Medical Image Analysis*, 2022, 75: 102277.

[2] Vrtovec T, Pernuš F, Likar B. A review of methods for quantitative evaluation of spinal curvature[J]. *European Spine Journal*, 2009, 18(5): 593-607.

[3] Yang G S, Fu X L, Xu N F, et al. A landmark estimation and correction network for automated measurement of sagittal spinal parameters[M]//Yang H, Pasupa K, Leung A C S, et al. *Neural information processing. Communications in computer and information science*. Cham: Springer, 2020, 1332: 213-221.

[4] Khanal B, Dahal L, Adhikari P, et al. Automatic Cobb angle detection using vertebra detector and vertebra corners regression[M]//Cai Y, Wang L, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 81-87.

[5] Zhang K L, Xu N F, Yang G S, et al. An automated Cobb angle estimation method using convolutional neural network with area limitation[M]//Shen D, Liu T, Peters T M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2019. Lecture notes in computer science*. Cham: Springer, 2019, 11769: 775-783.

[6] Chen K L, Peng C, Li Y, et al. Accurate automated keypoint detections for spinal curvature estimation[M]//Cai Y, Wang L, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 63-68.

[7] Wu H B, Bailey C, Rasoulinejad P, et al. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet[M]//Descoteaux M, Maier-Hein L, Franz A, et al. *Medical image computing and computer-assisted intervention-MICCAI 2017. Lecture notes in computer science*. Cham: Springer, 2017, 10433: 127-135.

[8] Sun H L, Zhen X T, Bailey C, et al. Direct estimation of spinal Cobb angles by structured multi-output regression [M]//Niethammer M, Styner M, Aylward S, et al. *Information processing in medical imaging. Lecture notes in computer science*. Cham: Springer, 2017, 10265: 529-540.

[9] Yi J R, Wu P X, Huang Q Y, et al. Vertebra-focused landmark detection for scoliosis assessment[C]//2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), April 3-7, 2020, Iowa City, IA, USA. New York: IEEE Press, 2020: 736-740.

[10] Guo Y, Li Y M, Zhou X W, et al. A keypoint transformer to discover spine structure for Cobb angle estimation[C]//2021 IEEE International Conference on Multimedia and Expo (ICME), July 5-9, 2021, Shenzhen, China. New York: IEEE Press, 2021.

[11] Ao Y Y, Wu H. Feature aggregation and refinement network for 2D anatomical landmark detection[J]. *Journal of Digital Imaging*, 2023, 36(2): 547-561.

[12] Reddy P K, Kanakatte A, Gubbi J, et al. Anatomical landmark detection using deep appearance-context network[C]//2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), November 1-5, 2021, Mexico. New York: IEEE Press, 2021: 3569-3572.

[13] Zhang C, Wang J, He J, et al. Automated vertebral landmarks and spinal curvature estimation using non-

- directional part affinity fields[J]. *Neurocomputing*, 2021, 438: 280-289.
- [14] Wang J S, Ma Y C, Huang S F, et al. A keypoint-based global association network for lane detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 1382-1391.
- [15] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 764-773.
- [16] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge[J]. *International Journal of Computer Vision*, 2015, 115(3): 211-252.
- [17] Wu H B, Bailey C, Rasoulinejad P, et al. Automated comprehensive Adolescent Idiopathic Scoliosis assessment using MVC-Net[J]. *Medical Image Analysis*, 2018, 48: 1-11.
- [18] Wang L S, Xie C, Lin Y, et al. Evaluation and comparison of accurate automated spinal curvature estimation algorithms with spinal anterior-posterior X-Ray images: the AASCE2019 challenge[J]. *Medical Image Analysis*, 2021, 72: 102115.