

基于嵌入注意力机制的 X 光脊椎骨角点定位模型

陈瑶, 余文俊*, 高永彬

上海工程技术大学电子电气工程学院, 上海 201620

摘要 脊柱侧弯是当今社会中常见的脊柱疾病, 在 X 光图像上快速而准确地定位脊椎骨角点并计算其 Cobb 角度数是医生诊断脊柱弯曲程度的金指标。针对 X 光骨科图片中其他器官的遮挡以及复杂背景干扰等问题, 提出一种基于嵌入注意力机制和向量损失模块的神经网络模型。所提模型以 vertebra-focused landmark detection (VFLD) 网络为基础网络, 在编码器和解码器之间嵌入旋转注意力机制模块加强网络对于脊椎骨深层、高维特征的提取, 抑制其他器官的干扰, 同时利用向量相似性的损失函数对网络进行训练。实验结果表明, 在 MICCAI 2019 公开脊椎挑战赛数据集中, 所提模型的对称平均绝对百分比误差准确度高达 9.31, 可以有效提高原模型检测椎骨角点能力。与现有的诸多模型相比, 其具有较高的准确率和稳健性。

关键词 图像处理; X 光图像; 神经网络; 注意力机制; Cobb 角

中图分类号 TP391.4 **文献标志码** A

DOI: 10.3788/LOP220851

X-Ray Spine Corner Localization using an Embedded Attention Mechanism-Based Model

Chen Yao, Yu Wenjun*, Gao Yongbin

School of Electronic and Electrical Engineering, Shanghai University of Engineering Science,
Shanghai 201620, China

Abstract Scoliosis is a common spinal disease in the current society. Therefore, it is important for doctors to diagnose the degree of spinal curvature to quickly and accurately locate the spinal bone corners on X-ray images and calculate their Cobb angles. In light of the occlusion of other organs and complex background interference in orthopaedic X-ray images, a neural network model based on the embedded attention mechanism, vector loss module, and vertebra-focused landmark detection (VFLD) network is proposed. The rotary attention mechanism module is embedded between the encoder and decoder to enhance the network's extraction of the deep and high-dimensional features of the spine bone, inhibit the interference of other organs, and allow the use of the vector similarity loss function to train the network. The experimental results show that the accuracy of the symmetrical mean absolute percentage error of the proposed model in the MICCAI 2019 open spine challenge dataset is as high as 9.31, and can effectively improve the ability of the original model to detect vertebral corners. Compared with many existing models, the proposed model has a higher accuracy and robustness.

Key words image processing; X-ray image; neural network; attention mechanism; Cobb angle

1 引言

脊椎骨侧弯包括结构性横移、旋转等, 严重可导致生活不便, 青少年特发性脊柱侧弯是一个人在青少年时期最常见的椎骨畸形疾病^[1]。目前, 衡量脊椎骨侧弯程度的金标准是 Cobb 角的度数, Cobb 角测量的准确性和鲁棒性取决于是否准确地标定出每一块脊椎骨的角点位置^[2]。现有的方法主要是临床经验丰富的骨

科医生通过手动标点进行 Cobb 角测量计算的, 这种方法不仅耗时费力, 而且主观性极大。即使是资历丰富的医生, 不同的医生测量出来的结果也可能相差很大。当患者数量增多、工作量增大时, 即使是同一名医生所测量的结果也会产生误差。因此, 开发一种具有准确性和鲁棒性的脊椎骨角点定位方法成为很多人的研究方向。

近年来, 许多人从不同方向对脊椎骨角点定位进

收稿日期: 2022-03-01; 修回日期: 2022-03-21; 录用日期: 2022-05-10; 网络首发日期: 2022-05-20

通信作者: *yuwenjun@sues.edu.cn

行研究。目前,椎骨角点定位研究主要分为传统方法和深度学习方法这两类。在传统方法中,Wills等^[3]等使用模板匹配的方法设计了包含4类大小不同的共56种模板,每类模板旋转14个角度,每次对同一个椎骨进行多次模板匹配,选择得分最大模板的4个顶点作为当前椎骨的顶点坐标,不过这种方法特别依赖于模板的设计,没有任何椎骨特征信息,泛化性很差。为了提取椎骨的特征信息,Zhen等^[4]提出了支持向量机回归方法,它使用基于监督学习的方法手工设计算法提取特征点,然后使用手工提取的特征拟合回归每个脊椎骨的角点,最终得到Cobb角。虽然与模板匹配的方法相比,支持向量机的方法增加了椎骨的特征信息,但是手工设计的特征提取方法不能适应于图片模糊的情况。随着深度学习在计算机领域的快速发展^[5-8],卷积神经网络(CNN)在医学X光图像领域取得了显著的成果^[9-10]。与基于模板匹配和支持向量机的传统方法相比,基于卷积神经网络的方法在准确性和鲁棒性上都有很大的提升,这些工作大多数将椎骨角点定位转化为椎骨识别任务。Chen等^[11]使用多边形回归模型与卷积神经网络混合训练的方法提取更加稳健的椎骨特征,在准确性和稳定性上比传统手工设计椎骨特征方法有了很大的提升,但是在遇到其他器官干扰时,会出现椎骨定位偏移的现象。Yang等^[12]通过深度监督卷积神经网络来预测每个椎骨质心,并将每个椎骨的质心输出到一张概率图上,根据质心坐标连线的斜率计算Cobb角度数。将所有质心输出到一张概率图上,可以在清晰的图片上抑制其他器官的干扰,但是在模糊图片上仍会出现预测偏移现象。为了在模糊图片上提取更多的特征,Liao等^[13]使用全卷积神经网络(FCN)和循环卷积神经网络(RNN)直接回归每个椎骨的质心坐标,根据质心坐标连线的斜率计算Cobb角度数,增加循环卷积神经网络之后,网络获得更多的特征,可抑制预测偏移的现象,但是并没有解决连续椎骨预测这一遗漏问题。BoostNet^[14]使用卷积神经网络提取整张图片的特征,直接回归得到每个顶点的位置,然后计算Cobb角度数,但回归网络中参数众多,医学数据集普遍存在数据量较少的问题,网络学习到的内容不充分。vertebra-focused landmark detection (VFLD)^[15]神经网络借鉴Liao等提出的网络和BoostNet,首先使用一张热力图取代回归方法,得到椎骨中心点坐标,然后依据中心点到4个角点的偏移量得到椎骨角点坐标。上述方法在较大程度上为椎骨角点定位奠定了基础,提高了角点定位网络的性能,但是在其他器官干扰严重、图片模糊不清的情况下,其精确性和鲁棒性需要进一步提高。

为了解决以上问题,本文以VFLD网络为基础框架,嵌入旋转注意力机制^[16]使得网络可以提取椎骨相互之间的空间关系,提高网络对于高维特征的关注,并且使用向量损失函数降低网络在模糊条件下的错误

推理。VFLD网络使用ResNet神经网络作为编码阶段的基础网络,可以在模糊图片上提取充足特征;网络输出中的一个分支为椎骨中心点热力图,将所有的椎骨中心点输出到一张热力图上,可以抑制其他器官干扰;网络不涉及回归的方法,较少的参数适用医学数据集数据量小的特点。但是VFLD神经网络存在细节信息缺失、椎骨连续性约束不强、各个椎骨分开对待缺乏一致性问题,本文针对这些缺点,提出一种融合旋转注意力机制、向量损失函数的脊椎骨X光角点定位网络模型。在MICCAI 2019公开脊椎挑战赛数据集的实验结果表明,所提模型的性能优于当前主流模型,展现出较好的准确性和鲁棒性。

2 脊椎骨X光图片角点定位模型

2.1 基本原理

所提模型以VFLD网络作为脊椎骨角点定位的基础网络,如图1(a)所示,主要包括编码部分、解码部分和输出部分。编码部分(字母E表示的部分)由ResNet的前5层卷积层构成,用于提取图像中的空间特征;解码部分(字母D表示的部分)为反向ResNet,根据编码阶段提取的特征值进行特征放大;输出部分包括3个子部分,第1个子部分是用来预测脊椎骨中心的热点图的,第2个子部分是预测脊椎骨中心点到4个角点的向量的,第3个子部分是中心点放缩回原始尺寸的偏移向量。在编码器和解码器之间的跳接层加入旋转注意力机制,对网络提取图片的特征进行优化,使得网络可以学习到脊椎骨之间的序列性关系和空间之间的联系,可以有效抑制其他器官的干扰,更加注意到脊椎骨的特征。在训练阶段,采用平移拼接的方法进行数据增强,提高网络的泛化能力。在实验过程中发现存在部分预测出现移位偏差的现象,对此提出向量损失函数,用来提高模型的精准性和鲁棒性。

2.2 旋转注意力机制

VFLD网络在编码部分和解码部分仅通过直接跳接层进行连接,跳接层缺乏有效的注意力机制,使得网络不能关注到脊椎骨两两相互之间、空间之间存在的关联特征。因此,在编码器和解码器之间嵌入旋转注意力机制对网络进行优化。旋转注意力机制由两个通道构成,能够在通道和空间这两个方向对脊椎骨的相互特征进行压缩与重新加权,提高网络对于重要特征的关注度,抑制其他器官信息的干扰。用图1(b)结构替换掉图1(a)中的跳接层结构(灰色箭头部分),得到改进的VFLD模型。图1(b)由两条支线组成,上面的旋转注意力支线用来提取空间和通道上的交叉关联信息,下面的平行支线用来提取空间信息。

在每一个分支的pool模块使用最大池化和平均池化,通过pool层可以将特征图从C维通道的 $\mathbf{z} \in \mathbf{R}^{C \times H \times W}$ 转化为两维通道的 $\mathbf{z} \in \mathbf{R}^{2 \times H \times W}$,使用两种不同的池化层可以提取更加丰富的特征信息;在Conv

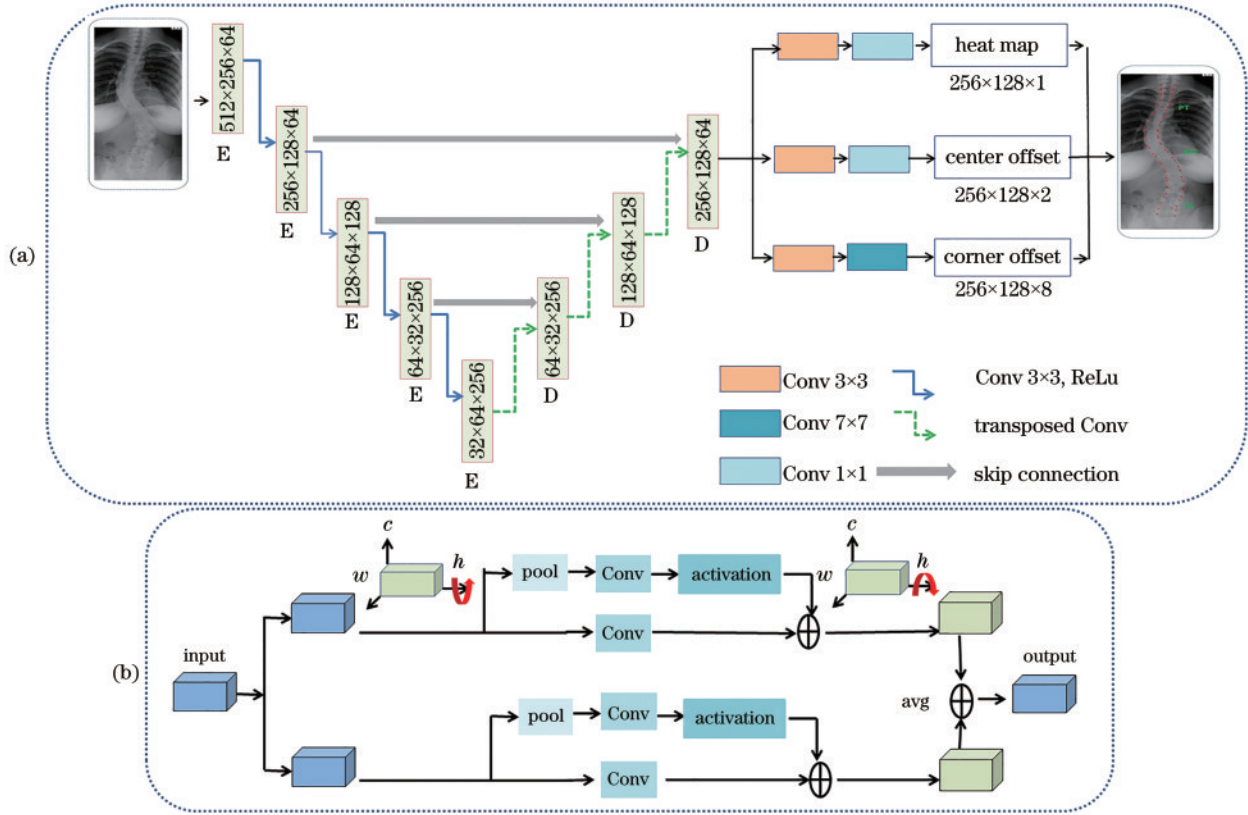


图 1 改进的 VFLD 网络结构。(a) VFLD 基础网络结构图; (b) 旋转注意力机制结构图

Fig. 1 Architecture of modified VFLD. (a) Basic network structure of VFLD; (b) architecture of rotational attention mechanism

模块中,通过卷积将两维通道的特征图 $z \in \mathbf{R}^{2 \times H \times W}$ 转化为 C 维通道的特征图 $z \in \mathbf{R}^{1 \times H \times W}$; activation 模块使用 Sigmoid 激活函数。

在上分支中,在通道和空间上建立注意力机制。将特征图沿着 h 轴逆时针旋转 90° , 旋转后的特征图为 $Z_0 \in \mathbf{R}^{W \times H \times C}$ 。首先将 Z_0 传递给 pool 层, 经过 pool 层的平均池化和最大值池化之后, 特征图 Z_0 大小变为 $Z_0 \in \mathbf{R}^{2 \times H \times C}$; 然后将特征图 Z_0 输入 Conv 模块, Conv 模块使用 $k \times k$ 的卷积核, Z_0 经过 Conv 模块之后输出的特征图大小为 $Z_0 \in \mathbf{R}^{1 \times H \times C}$; 经过 activation 模块后与原特征图相加, 然后顺时针旋转 90° 。传统注意力机制仅在空间维度增加注意力模块, 从而缺失了通道维度的信息, 所提模型通过上分支注意力机制在特征块的通道 C 和高度 W 建立联系, 让网络学习到 C 维和 W 维的交叉特征。可理解为, 将一张宽、高、通道数分别为 w, h, c 的直立 CT 椎骨图片放平, 用宽、通道维度的二维卷积取代原本在宽、高维度的二维卷积, 将通道信息补充到网络中来, 让网络学习到相邻椎骨之间的关联信息。

在下分支中, 在空间上建立注意力机制。将特征图 Z_1 直接传递给 pool 模块, Z_1 的大小从 $Z_1 \in \mathbf{R}^{C \times W \times H}$ 变成 $Z_1 \in \mathbf{R}^{2 \times W \times H}$; 然后将特征图 Z_1 输入 Conv 模块, Conv 模块使用 $k \times k$ 的卷积核, Z_1 经过 Conv 模块之后

输出的特征图大小为 $Z_1 \in \mathbf{R}^{1 \times W \times H}$; 经过 activation 模块后与原特征图相加, 最终将上下两个支线的特征图进行平均求和。

2.3 损失函数

本网络框架的损失函数主要包括 4 个部分, 分别是中心点损失函数、中心点偏移损失函数、脊椎骨中心点到 4 个顶点的偏移向量损失函数和在实验中根据实验情况添加的向量损失函数。

在实验中发现, 当椎骨弯曲的角度过大或者图片模糊性较大时, 网络的预测输出会出现偏差, 如图 2 左图所示, 网络在预测腰部椎骨位置时, 图片模糊性较大, 导致出现将其他器官预测成椎骨的错误现象(箭头指向部位)。经过多次实验总结规律得出: VFLD 网络

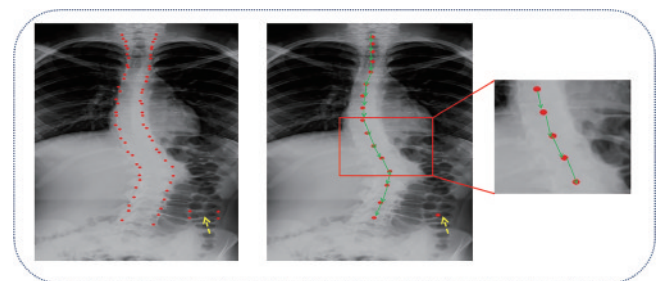


图 2 向量损失模块示意图

Fig. 2 Schematic diagram of vector loss module

在输出部分将所有椎骨中心点输出到一张热力图上,通过一张热力图让网络学习到椎骨之间的相互关系,直接约束性不明显,缺少上下两个相邻椎骨之间的限制关系。因此,提出向量损失模块解决这个问题。在训练模型时,网络的输出多一项相邻脊椎骨之间的向量值,如图 2 右图所示,每个椎骨都与其相邻的椎骨有一个方向向量,如矩形框所示,每个向量以上一个椎骨中心为起点、下一个椎骨中心为终点构成一个方向向量,本实验使用方向向量余弦值来限制出现错误预测的情况。

加上向量损失模块后,当模型遇到弯曲角度过大、X 光图片模糊或者其他器官相似性较大的情况时,向量损失便会根据前后椎骨之间的关系纠正预测的位置,将预测偏差或者遗漏的现象去除掉,从而提高模型在高模糊、弯曲角度过大以及 X 光脊椎骨图像上的准确性和鲁棒性。向量损失模块损失公式如下:

$$L_{\text{vector_loss}} = 1 - \cos(\mathbf{P}_{\text{vector}}, \mathbf{G}_{\text{t-vector}}), \quad (1)$$

式中: $\mathbf{P}_{\text{vector}}$ 表示网络框架预测的向量值; $\mathbf{G}_{\text{t-vector}}$ 表示相邻两个椎骨之间的真实向量值; \cos 表示使用余弦函数进行计算。

中心点损失函数是网络框架定位椎骨的第一步,它是基于关键点的热力图,常用在姿态估计上。对于每个椎骨的中心点,它的真实值是一个 2D 的高斯分布图,分布形式可以用 $\exp\left[-\left(\frac{x^2 + y^2}{2\sigma^2}\right)\right]$ 来表示,其中 σ 由 X 光图片的尺寸决定。使用标准的 Focal 损失^[17]来优化这些参数, Focal 损失具体形式如下:

$$L_{\text{hm}} = -\frac{1}{N} \sum_i \begin{cases} (1 - p_i)^\alpha \ln p_i, & y_i = 1 \\ (1 - y_i)^\beta (p_i)^\alpha \ln(1 - p_i), & \text{otherwise} \end{cases}, \quad (2)$$

式中: i 代表每一块椎骨的编号; N 代表总共的位置数量; p 表示网络预测的值; y 表示真实值; α 和 β 为超参数,分别设置为 2、4。

中心点损失函数是将热力图放大到原始图片大小时的偏移量,因为分辨率的不同会导致产生小数点后的数字,简单使用四舍五入不能解决小数点后的部分与整数之间的差异,因此本实验使用中心点偏移损失来解决这个问题。真实的中心点坐标与预测偏移量的差异值为

$$\varphi = \left(\frac{c_x}{s} - \left\lfloor \frac{c_x}{s} \right\rfloor, \frac{c_y}{s} - \left\lfloor \frac{c_y}{s} \right\rfloor \right), \quad (3)$$

式中: c_x 、 c_y 为放大前中心点坐标; s 为实际坐标; $\lfloor \cdot \rfloor$ 为向下取整。使用 L1 损失来优化这个偏移量。

中心点到 4 个顶点的偏移向量使用 L1 损失进行优化。L1 损失公式如下:

$$L_{1\text{ loss}} = \frac{\sum_{i=1}^n |y_i - y_i^p|}{n}, \quad (4)$$

式中: y 和 y^p 分别代表真实向量和神经网络预测的向量; i 表示椎骨的序号; n 为 17, 代表人体的 17 块脊椎骨骼。

2.4 数据增强

针对医学数据集数据量不大,而神经网络又需要大量数据进行训练的情况,本实验在训练之前对数据集进行扩充。普通的数据增强策略如旋转、仿射变换等对医学数据集不适应,会扰乱 X 光图片所携带的信息,本实验使用平移拼接的策略对数据进行扩充,并且只使用左右平移的策略进行裁剪,如图 3 所示。这样扩充数据集,既不影响脊椎骨骼之间的相互关系,也可以增强神经网络模型的泛化能力。

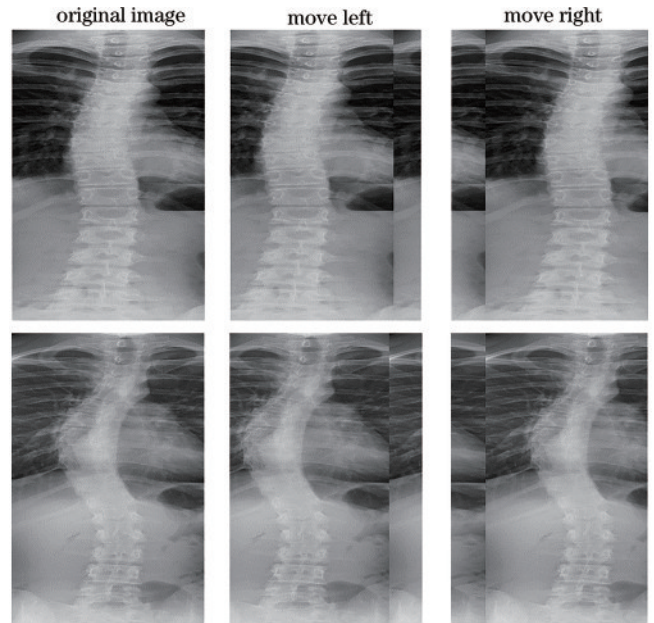


图 3 数据增强效果图

Fig. 3 Design sketch of data enhancement

3 实验分析与讨论

3.1 数据集

所采用的数据集为 AASCE MICCAI 2019 脊椎挑战赛的公开数据集,包括 609 张训练数据集、98 张测试数据集。每张图片都经过了脱敏处理,去除了 X 光图片所携带的患者姓名、年龄信息。每张图片都是统一的 755×2125 的长宽尺寸,相对应的标签是临床经验丰富的骨科医生手工标注的。实验中,将训练数据集按照 9:1 的比例进行切分,用于训练集和验证集对神经网络模型进行训练。

除此之外,使用长征医院骨科颈椎骨数据集,一共包括 791 张训练集和 80 张测试集,每张颈椎骨数据都经过两名医生手工标记,实验中将训练集按照 9:1 的比例进行切分,用作训练集和验证集。

3.2 实验配置

本实验是在 Ubuntu18.04 系统下进行的,神经网络框架是在 PyTorch1.8 环境下搭建的,显卡使用 1080TI。batch size 前期设置为 8、后期设置为 4,权重的学习率设置为 1.5×10^{-4} ,使用预训练的 ResNet,训练的轮数设置为 100。将原始图像放缩为 1024×512 后,输入进神经网络。

3.3 评价指标

使用与 AASCE MICCAI2019 挑战赛相同的评价指标对称平均绝对百分比误差(SMAPE)来评价神经网络:

$$E_{\text{SMAPE}} = \frac{1}{N} \sum_{j=1}^M \frac{\sum_{i=1}^3 (|p_{ji} - g_{ji}|)}{(p_{ji} + g_{ji})}, \quad (5)$$

式中: i 代表网络预测不同部位的 Cobb角; j 代表第 j 张

图片; N 代表所有的测试图片数量之和; p 和 g 分别代表预测的 Cobb角和真实的 Cobb角。AASCE MICCAI 2019 脊椎挑战赛只用一个主体 SMAPE来评价神经网络预测的好坏,本实验还对比了上胸端(PT)、主胸端(MT)、近腰端(TL)等 3 个不同部位的 SMAPE 值。

使用平均检测误差来评价网络框架对于椎骨角点定位的准确度,平均检测误差定义如下:

$$E_{\text{dec}} = \frac{1}{M} \sum_{i=1}^M \|d_i - g_i\|^2, \quad (6)$$

式中: d 代表网络预测的角点坐标; g 代表实际的角点坐标值; M 代表总共的测试集图片数量。

3.4 结果分析

为了检验所提网络的有效性,将原始网络与改进的网络进行对比实验,结果如表 1 所示。

表 1 模型改进前后的性能指标

Table 1 Comparison indicators of model before and after improvement

| Method | SMAPE / % | SMAPE _{PT} / % | SMAPE _{MT} / % | SMAPE _{TL} / % | E_{dec} / pixel | FPS | Weight / MB |
|------------------|-----------|-------------------------|-------------------------|-------------------------|--------------------------|-------|-------------|
| Original network | 10.81 | 6.26 | 18.04 | 23.42 | 50.11 | 11.40 | 97 |
| Proposed network | 9.31 | 5.75 | 15.18 | 22.35 | 48.75 | 11.34 | 98.5 |

相比于原始网络,改进模型的整体 SMAPE 值从 10.81% 提升到了 9.31%,精度值提升了 1.50 个百分点。在上胸端,改进模型的 SMAPE 值较原始网络提升了 0.51 个百分点;在主胸端,因为无明显其他器官遮挡,椎骨图像的清晰度明显,改进模型的 SMAPE 值较原始网络提升了 2.86 个百分点;在近腰端,其他器官遮挡较多,所提模型的 SMAPE 值较原始网络提升了 1.07 个百分点。所提模型的平均检测误差比原始网络提升了 1.36 个像素。网络模型权重的大小对比

之前仅增加了 1.5 MB,同时网络的运行速度并未减慢。综上,改进后的网络模型不仅在整体的 SMAPE 值上有提高,在不同部位的 SMAPE 值上均有提高,表明改进的网络模型是有效的。

为了验证改进模型的先进性,进行了性能测试,并与当前主流的算法以及 MICCAI2019 挑战赛的算法进行对比,如表 2 所示。在 MICCAI2019 脊椎骨挑战赛数据集上,所提模型在 SMAPE 值上达到最高。

表 2 MICCAI 2019 脊椎挑战赛数据集结果对比

Table 2 Result comparison of MICCAI 2019 spine challenge dataset

unit: %

| Method | SMAPE | SMAPE _{PT} | SMAPE _{MT} | SMAPE _{TL} |
|-------------------------------------|-------------|---------------------|---------------------|---------------------|
| Original method | 10.81 | 6.26 | 18.04 | 23.42 |
| Seg4Reg ^[18] | 22.25 | | | |
| Regression method ^[19] | 23.43 | 16.38 | 30.27 | 35.61 |
| ACAD ^[20] | 25.69 | | | |
| Segmentation method ^[21] | 16.48 | 9.71 | 25.97 | 33.01 |
| SCG-Net ^[22] | 22.17 | | | |
| AAKD ^[23] | 22.16 | | | |
| Proposed method | 9.31 | 5.75 | 15.18 | 22.35 |

与传统的基于回归的方法和基于分割的方法相比,所提模型的整体 SMAPE 值分别提升了 14.12 个百分点和 7.17 个百分点。与 MICCAI 2019 挑战赛系列算法相比,所提模型在准确性方面对比最优的 AAKD 网络框架有 12.85 个百分点的提高,表现优异。综上,改进的网络模型具有良好的先进性。图 4 展示了改进模型的效果图。从第 1 行左边的案例可以看

出,即使在椎骨弯曲角度很大的情形下,网络框架也能准确地定位到椎骨角点的位置;从第 1 行右边的案例可以看出,脊椎骨的整个 CT 影像十分模糊,椎骨的边界线与其相邻的背景器官区分性较小,但所提网络框架依然可以准确定位到椎骨角点;第 2 行左边案例集弯曲角度大、图像模糊为一体,在腰部位置放大后也很难分辨椎骨的具体边界,需要临床经验丰富的骨科医

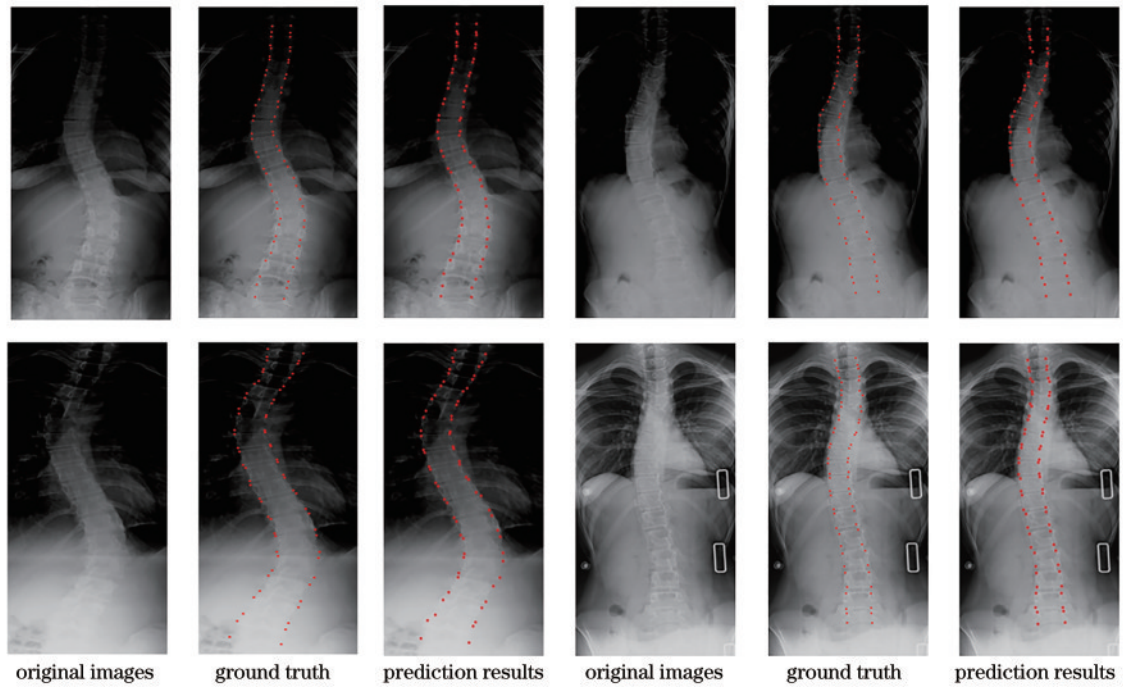


图 4 识别结果图

Fig. 4 Identification result diagrams

生才能准确定位椎骨角点位置,所提网络也能准确预测椎骨角点位置;第 2 行右边案例有明显的外界干扰物,所提网络也可以给出准确的预测结果。综上,在 X 光图片在十分模糊、弯曲角度特别大的情况下,所提网络也可以得到准确的脊椎骨角点定位。

为了验证改进模型的泛化性,使用长征医院骨科颈椎骨数据集对网络进行泛化能力测试,结果如表 3 所示。从表 3 可以看出:改进模型的整体 SMAPE 值从 8.44% 提升到了 6.21%,精度值提升了 2.23 个百分点;改进模型的平均检测误差从 16.41 个像素提升到了 15.21 个像素,减小了 1.2 个像素;在计算速度上并未因为增加模块而变慢;权重大小对比基础网络仅增加了 1.5 MB。综上,改进后的网络模型不仅可以应用于脊椎骨角点定位,还适用于颈椎骨角点定位,即可以适用于不同的角点检测情景。图 5 展示了所提网络框架对于颈椎骨 X 光图片的预测结果,可以看出所提网络可以准确预测每个颈椎骨角点的位置,说明所提网络模型具有泛化性。

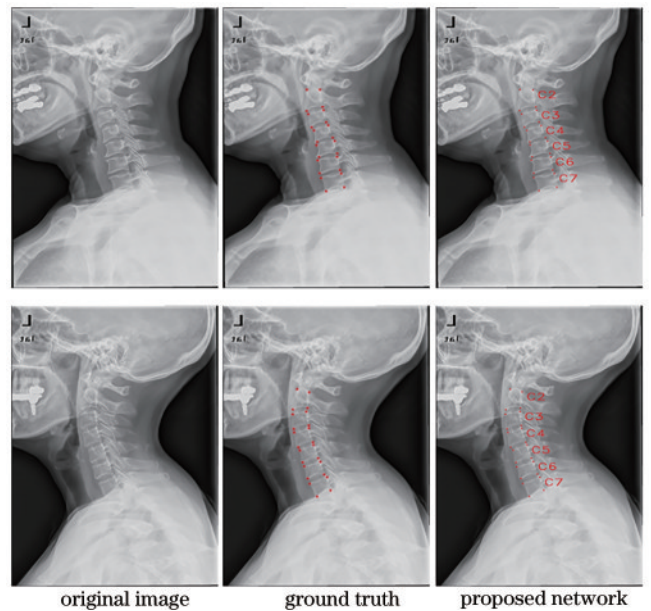


图 5 颈椎骨识别结果图

Fig. 5 Recognition result images of cervical vertebra

表 3 颈椎骨数据集改进前后对比实验
Table 3 Comparison experiment before and after improvement of cervical vertebrae dataset

| Method | SMAPE /% | E_{dec} /pixel | FPS | Weight /MB |
|-----------------|----------|------------------|------|------------|
| Original method | 8.44 | 16.41 | 14.6 | 97 |
| Proposed method | 6.21 | 15.21 | 14.6 | 98.5M |

3.5 消融实验

为了进一步研究探讨所提网络对于网络性能的影响,进行了消融实验对此进行验证,结果如表 4 所示。从表 4 可以看出,加入旋转注意力机制之后的模型在 SMAPE 值上提升了 1.39 个百分点。可归因于注意力机制模块提升网络对于两两相邻的脊椎骨之间特征的捕捉,提升了对于脊椎骨角点特征的捕捉,使得网络框架可以抑制其他器官干扰,更加精准地预测脊椎骨角点的位置。加上向量损失模块之后, SMAPE 值也有一

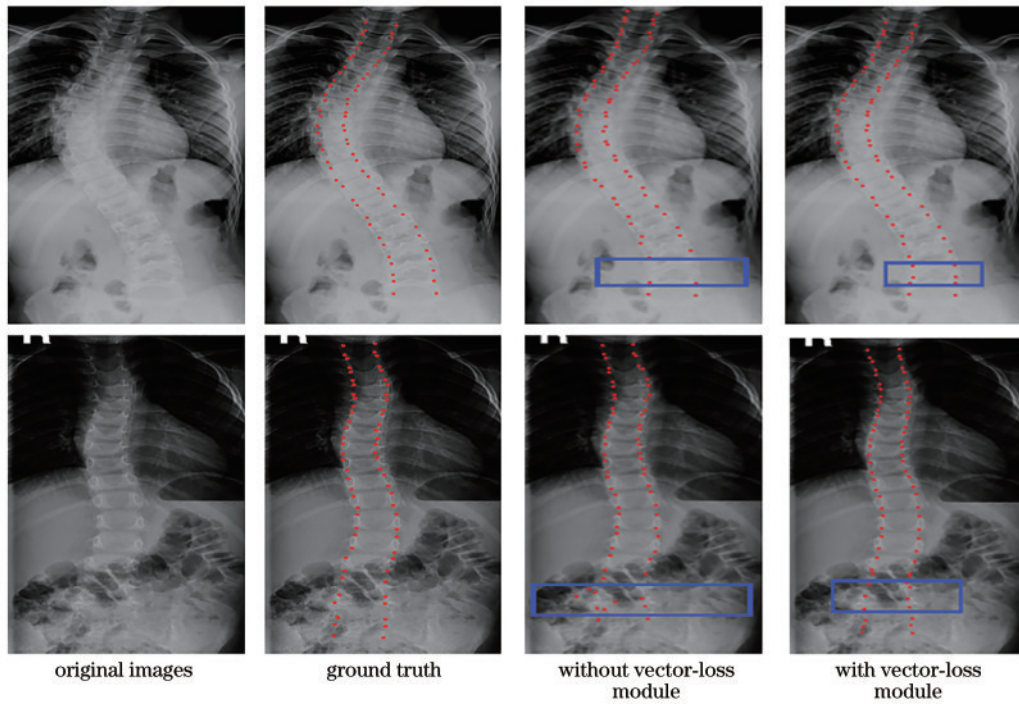


图 6 消融实验

Fig. 6 Ablation experiment

表 4 消融实验

Table 4 Ablation experiment

| Method | SMAPE / % |
|---|-----------|
| Original network | 10.81 |
| Original network + rotational attention module | 9.42 |
| Original network + rotational Attention module + vector loss module | 9.31 |

定的提升,不过提升的幅度不大,归因于损失函数对网络带来的影响小于改进网络结构所带来的影响。图 6 中红点表示椎骨的角点位置坐标,可以看出向量损失模块可以在网络预测错误时进行及时矫正,对于图片模糊地方,向量损失模块可以很好地进行纠正。从图 6 第 1 行图片可以看出,未加向量损失模块的网络框架出现了连续椎骨遗漏的现象(蓝色方框部位),而加上向量损失模块之后这个问题被有效解决;第 2 行图片中,在模糊的腰腹部加上向量损失模块的网络框架可以将预测偏移的椎骨(蓝色方框部位)拉回正确的位置。根据表 3 和图 6 可以得出:旋转注意力机制模块和向量损失模块的使用对结果起到了提升作用,而且主要的提升来自于注意力机制模块。

4 结 论

基于 X 光 CT 医学影像的脊椎骨角点定位自动化、智能化实现的任务,以 VFLD 网络作为基础网络模型,融合旋转注意力机制,使用数据增强策略扩充

数据,并根据实验中出现的状况提出了向量损失模块对基线网络模型进行优化,解决了在模糊 CT 医学影像上椎骨角点定位不准的问题。旋转注意力机制可以使得网络忽略其他器官的干扰,专注于捕捉脊椎骨角点的位置特征,增加脊椎骨相互之间的特征;数据增强策略可以让网络框架更加具有泛化性和鲁棒性;向量损失模块使得网络在图片十分模糊、其他器官干扰严重的情况下可以正确预测角点位置。在 MICCAI2019 脊椎挑战赛数据集上,改进的网络模型优于当前其他的椎骨模型,具有较好的准确性和鲁棒性。虽然在网络的识别效果上已经达到应用的要求,但是在模型应用上考虑到各个基层医院的科研经费,落地时使用轻量化的模型可以减少仪器的成本费用,更加容易推广,未来研究重点是使用新颖的剪枝技术、模型压缩技术等对所提模型进行压缩,以便更好地移植于手持设备。

参 考 文 献

- [1] Carr A J. Adolescent idiopathic scoliosis in identical twins [J]. The Journal of Bone and Joint Surgery. British Volume, 1990, 72(6): 1077.
- [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] Wills B P D, Auerbach J D, Zhu X W, et al. Comparison of Cobb angle measurement of scoliosis radiographs with preselected end vertebrae: traditional versus digital acquisition[J]. Spine, 2007, 32(1): 98-105.
- [4] Zhen X T, Wang Z J, Islam A, et al. Multi-scale deep

- networks and regression forests for direct bi-ventricular volume estimation[J]. *Medical Image Analysis*, 2016, 30: 120-129.
- [5] 张珂, 张良. 复杂背景下多尺度 X 光违禁品检测[J]. *激光与光电子学进展*, 2021, 58(22): 2210002.
Zhang K, Zhang L. Multi-scale detection for X-ray prohibited items in complex background[J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 2210002.
- [6] 郭守向, 张良. Yolo-C: 基于单阶段网络的 X 光图像违禁品检测[J]. *激光与光电子学进展*, 2021, 58(8): 0810003.
Guo S X, Zhang L. Yolo-C: one-stage network for prohibited items detection within X-ray images[J]. *Laser & Optoelectronics Progress*, 2021, 58(8): 0810003.
- [7] 郭瑞鸿, 张莉, 杨莹, 等. 基于改进 SSD 的 X 光图像管制刀具检测与识别[J]. *激光与光电子学进展*, 2021, 58(4): 0404001.
Guo R H, Zhang L, Yang Y, et al. X-ray image controlled knife detection and recognition based on improved SSD[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0404001.
- [8] 胡均平, 宋菲菲, 罗春雷, 等. 一种基于自适应角点检测的非量测畸变校正方法[J]. *激光与光电子学进展*, 2021, 58(12): 1210012.
Hu J P, Song F F, Luo C L, et al. Non-metric distortion correction method based on adaptive corner detection[J]. *Laser & Optoelectronics Progress*, 2021, 58(12): 1210012.
- [9] 赵贺, 张金秀, 张正刚. 基于 NSCT 与 DWT 的 PCNN 医学图像融合[J]. *激光与光电子学进展*, 2021, 58(20): 2017002.
Zhao H, Zhang J X, Zhang Z G. PCNN medical image fusion based on NSCT and DWT[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 2017002.
- [10] 曹斌, 杨锋, 马金剛. 深度学习方法在肺结节诊断中的应用[J]. *激光与光电子学进展*, 2021, 58(16): 1600005.
Cao B, Yang F, Ma J G. Application of deep learning methods in diagnosis of lung nodules[J]. *Laser & Optoelectronics Progress*, 2021, 58(16): 1600005.
- [11] Chen H, Shen C Y, Qin J, et al. Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9343: 515-522.
- [12] Yang D, Xiong T, Xu D G, et al. Automatic vertebra labeling in large-scale 3D CT using deep image-to-image network with message passing and sparsity regularization [M]//Niethammer M, Styner M, Aylward S, et al. *Information processing in medical imaging. Lecture notes in computer science*. Cham: Springer, 2017, 10265: 633-644.
- [13] Liao H F, Mesfin A, Luo J B. Joint vertebrae identification and localization in spinal CT images by combining short- and long-range contextual information [J]. *IEEE Transactions on Medical Imaging*, 2018, 37(5): 1266-1275.
- [14] 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.
- [15] 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, April 3-7, 2020, Iowa City, IA, USA. New York: IEEE Press, 2020: 736-740.
- [16] Misra D, Nalamada T, Arasanipalai A U, et al. Rotate to attend: convolutional triplet attention module[C]//2021 IEEE Winter Conference on Applications of Computer Vision, January 3-8, 2021, Waikoloa, HI, USA. New York: IEEE Press, 2021: 3138-3147.
- [17] Law H, Deng J. CornerNet: detecting objects as paired keypoints[J]. *International Journal of Computer Vision*, 2020, 128(3): 642-656.
- [18] Lin Y, Zhou H Y, Ma K, et al. Seg4Reg networks for automated spinal curvature estimation[M]//Cai Y L, Wang L S, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 69-74.
- [19] Wang L S, Xu Q H, Leung S, et al. Accurate automated Cobb angles estimation using multi-view extrapolation net[J]. *Medical Image Analysis*, 2019, 58: 101542.
- [20] Khanal B, Dahal L, Adhikari P, et al. Automatic Cobb angle detection using vertebra detector and vertebra corners regression[M]//Cai Y L, Wang L S, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 81-87.
- [21] 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 and Mathematical Methods in Medicine*, 2019, 2019: 6357171.
- [22] Wang S X, Huang S H, Wang L S. Spinal curve guide network (SCG-net) for accurate automated spinal curvature estimation[M]//Cai Y L, Wang L S, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 107-112.
- [23] Chen K L, Peng C, Li Y, et al. Accurate automated keypoint detections for spinal curvature estimation[M]//Cai Y L, Wang L S, Audette M, et al. *Computational methods and clinical applications for spine imaging. Lecture notes in computer science*. Cham: Springer, 2020, 11963: 63-68.