

# 基于Transformer与Vector Loss模块的椎骨Cobb角点定位网络

陈瑶,高永彬,熊玉洁

上海工程技术大学电子电气工程学院,上海 201620

**【摘要】目的:**通过定位脊椎骨角点计算Cobb角度数。**方法:**使用神经网络的方法进行脊椎骨角点坐标的定位,通过嵌入Transformer与Vector Loss模块解决了在清晰度不高、拍摄角度不同的X光图像上计算Cobb角度数精确不高的问题。**结果:**在MICCAI 2019公开脊椎挑战赛数据集中,本文方法的平均对称百分比(SMAPE)高达9.01。相较于最新的方法,该方法在SMAPE值上提高了1.80。**结论:**本文所提出的算法嵌入Transformer与Vector Loss模块,与现有的诸多算法相比,具有较高的准确率和稳健性,可以辅助临床医生选择适合患者的治疗方案。

**【关键词】**X光图像;Cobb角;Transformer;Vector Loss;神经网络;辅助诊断

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2022)11-1393-08

## Vertebral Cobb corner localization using neural network with Transformer and Vector Loss modules

CHEN Yao, GAO Yongbin, XIONG Yujie

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

**Abstract: Objective** To calculate the Cobb angle by locating the vertebral corners. **Methods** Neural network was used to locate the coordinates of the vertebral corners. By embedding Transformer and Vector Loss modules, the problem of the poor accuracy in calculating the Cobb angle in X-ray images with low definition and different shooting angles was solved. **Results** In the MICCAI 2019 Open Spine Challenge data set, the proposed method achieved a symmetric mean absolute percentage error (SMAPE) as high as 9.01. Compared with the latest methods, the proposed method improved the SMAPE by 1.80. **Conclusion** The proposed algorithm with Transformer and Vector Loss modules is superior to the existing algorithms in accuracy and robustness, and can assist clinicians in selecting the treatment schemes suitable for patients.

**Keywords:** X-ray image; Cobb angle; Transformer; Vector Loss; neural network; auxiliary diagnosis

### 前言

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,椎骨侧弯包括结构性横向、旋转变形等,严重可导致生活不便<sup>[1]</sup>。在普通人群中,椎骨侧弯的发病率为2.5%,0.25%的患者需要进行手术治疗<sup>[2]</sup>。为临床医生提供早期的辅助诊疗可以限制疾病的发展,尽早发现脊柱侧弯的症状会对以后产生潜在的、积极的影响。早期脊椎侧弯的发现和尽早治疗可以减少或避免手术<sup>[3]</sup>。Cobb角是评价脊柱侧弯程度的金标准<sup>[4]</sup>。如图1所示,将一张X

光图像分为3个部分,分别为胸椎近端(PT)、主胸近端(MT)及腰椎近端(TL),选择相对于水平线倾斜最大的倾斜线进行夹角测量,测量所得到的角度便是Cobb角<sup>[5-6]</sup>。Cobb角测量的准确性取决于是否准确标记出每一块椎骨的角点位置。现存的方法需要临床经验丰富的骨科医生手动标点进行Cobb角计算,这种方法不仅费时费力而且存在很大的主观性,不同的骨科医生所测量出来的角度可能相差甚远。因此开发一种准确且快速的Cobb角测量方法成为了研究热点<sup>[7]</sup>。

Wills等<sup>[8]</sup>使用模板匹配方法设计4种大小的模板,每个模板旋转14个不同的角度,一共56个模板进行Cobb角计算。Prabhu等<sup>[9]</sup>和Anitha等<sup>[10]</sup>使用滤波与激活轮廓组合的方法对脊椎骨进行定位,以便从斜率推导出Cobb角。支持向量机回归算法<sup>[11]</sup>、随机森林回归算法、S2VR<sup>[12]</sup>使用基于监督学习的方法

**【收稿日期】**2022-01-15

**【基金项目】**上海市科委重点项目(18411952800)

**【作者简介】**陈瑶,硕士研究生,研究方向:计算机视觉、智慧医疗,E-mail: chenyaoyao9635@163.com

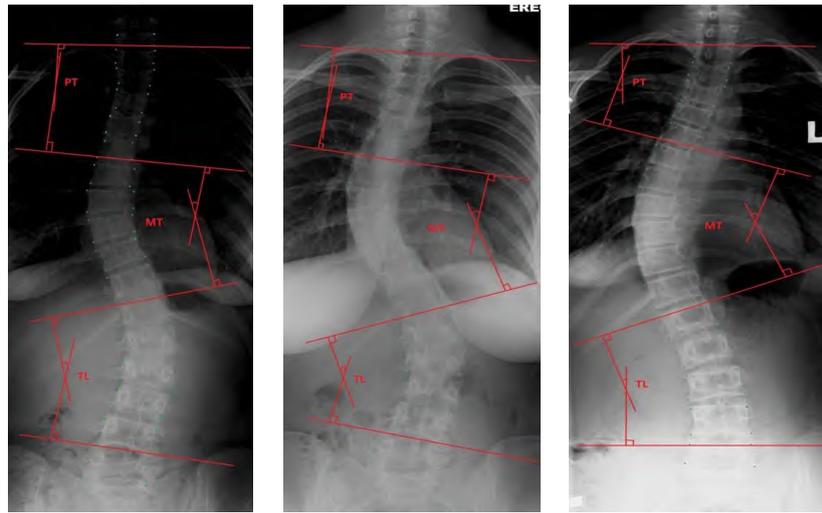


图1 基于X光图像的Cobb角计算

Figure 1 Calculation of Cobb angle in X-ray image

法,手工提取特征后,直接使用手工提取的特征拟合回归每个椎骨的顶点得到Cobb角。BoostNet通过卷积神经网络提取整张图像的特征,然后回归每个顶点的位置得到Cobb角<sup>[13]</sup>。Wang等<sup>[14]</sup>使用分割的方法对椎骨的Cobb角进行测量。这些方法都是基于U-Net<sup>[15]</sup>或V-Net<sup>[16]</sup>的改进版本,首先分割出每一块椎骨,然后确定4个顶点的位置,最后根据每个顶点的坐标计算出Cobb角的度数。

虽然已有的工作取得了不错的进展,但是还是存在一些问题。传统的算法中基于模板匹配的算法依赖于手工设计的模板长宽比例大小和旋转的角度,设计的模板有一些变化便会对结果产生很大的影响。基于数学方法的模型在计算上十分复杂且极易受到X射线变化而产生误差。深度学习中基于回归的方法可以利用提取图像的全局信息,但是回归点与高维特征之间需要密集的参数映射和大量的计算成本,预测的结果偏差较大;基于深度学习的分割方法需要计算并预测每个像素的类别,运算量跟图像的大小成正比,由于运算量的限制往往需要将原始图像调整为较小的尺寸,分割法只关注当前的区域是否为需要的对象,对于图像质量差的图像会将一块椎骨切分成两块的问题,对图像质量要求较高,难以有效分割附着度高、离散性大的椎骨。而且这些方法都是独立地预测每一块椎骨的位置,没有将椎骨所具有的两两之间的相互信息利用起来。为解决这个问题,Yi等<sup>[17]</sup>提出一种以检测椎骨中心关键点来进行椎骨检测的方法,将椎骨的中心点输出到一张蒙板上,希望通过一张蒙板隐性让网络学习椎骨互相之间的相互关系。但是,每个人的椎骨分布各不相同,通过这种隐性的约束让网络去学习椎

骨之间的相互关系,对于网络有很大的挑战性。

目前自动化椎骨Cobb角测量面临的问题有:(1)医学相关的影像数据稀缺,而神经网络需要大量数据进行训练,如果设计的网络过于浅显,则不能很好地学习图像的特征;如果网络设计过于复杂,在小量的数据集上会导致过拟合现象。(2)同一张X光图像中的每一块椎骨是两两之间相互联系的,前人使用的深度学习的方法都是将每一块椎骨隔离开,单独对待,缺少一致性。(3)图像的对比度较低,椎骨与周围的器官之间区分性不大,需要更加精确的网络框架进行特征提取。针对自动化椎骨Cobb角测量面临的问题,本研究融合神经网络中的目标检测以及特征点检测,提出一种嵌入Transformer机制的椎骨角点检测框架。主要贡献如下:(1)针对医学数据集较少的问题,本研究使用镜像与加噪声组合的方法进行数据扩充,这样扩充的数据不会改变原来的椎骨的相对位置。(2)针对椎骨之间相互缺少一致性的问题,本研究将Transformer模块融合到关键点检测网络中,以便于更好地提取高维张量中椎骨之间的相互关系,解决椎骨相互之间独立的问题。(3)提出一种向量损失,可以充分利用椎骨之间的位置空间关系解决对比度较低,椎骨与周边器官区分性不大的问题。

## 1 方法

本研究首先通过数据增强的方法扩充训练集,然后送入本研究提出的嵌入Transformer机制的神经网络框架,直接预测Cobb角的度数以及椎骨的坐标值。神经网络结构如图2所示,由编码器、解码器、Transformer模块组成。编码器使用ResNet网络,去

掉最后的全连接层, 将提取的特征块传递给 Transformer 模块; Transformer 模块将卷积神经网络提取的特征图展开成一维特征, 与位置编码结合后, 经过 Transformer 自带的编码器、解码器提取空间相

关特征, 最后输出一维特征, 将一维特征转成张量块后, 送入解码器; 解码器获取特征块后将查询编码阶段同等大小的特征块, 然后进行反卷积操作, 最终得到中心点坐标的热力图(Heatmap)以及偏移向量。

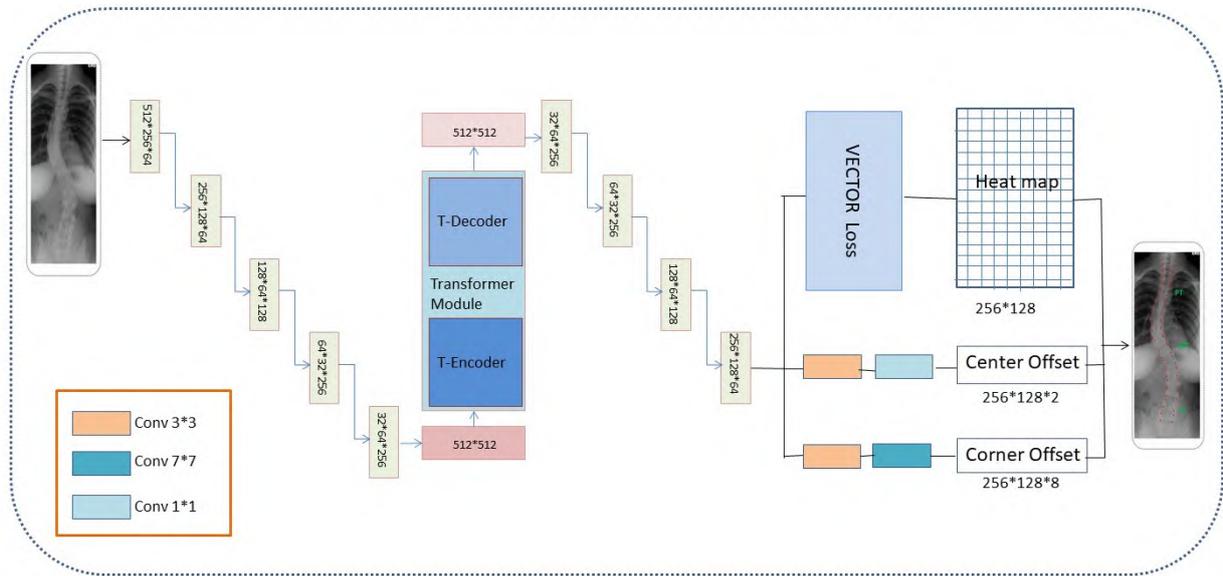


图 2 网络框架图

Figure 2 Network framework

### 1.1 数据增强

常用的图像增强处理有旋转、裁剪、仿射变换等, 但这些方法并不适应于本研究, 因为人体的脊椎骨的骨骼数量是一定的, 使用裁剪的方法只能裁剪出部分的骨骼, 图像不能携带整体骨骼的信息, 而仿射变换和旋转会导致脊椎骨骼的位置以及脊椎骨的大小发生变化, 对于网络预测椎骨位置以及定位 4 个角点的位置会产生反向影响。综上各种情况, 本研究采用图片镜像和加入噪声的方法产生新的图像, 更好地模拟原始的 X 光图像, 在扩充数据的同时更好地抑制过拟合。为了不破坏图像的原始结构, 本研究采用高斯噪声的方法, 使得像素值在一定范围内变化, 高斯噪声插入公式如下:

$$P_i^1 = P_i^0 + \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

其中,  $P_i^1$  为插入后的像素值;  $P_i^0$  为插入前的像素值;  $\mu$  为高斯分布的期望;  $\sigma$  为高斯分布的方差。通过镜像以及加入高斯噪声后的图像如图 3 所示。图像在增加噪声之后椎骨的相对位置并没有被破坏, 椎骨之间相互特征并没有被破坏。本研究在原数据集的基础之上将数据集扩充至原来数据集的 3 倍。

### 1.2 Transformer Encoder-Decoder 模块结构

Transformer Encoder-Decoder 模块如图 4 所示,

由两部分组成, 一个编码器和一个解码器。本研究使用 6 个头的多头 Transformer 模块, 循环编解码层数设置为 6。

编码器通过  $1 \times 1$  的卷积将多通道的特征图  $f$  从通道  $C$  转化为较低层的通道  $d$ , 得到一个新的特征图  $z_0 \in R^{d \times H \times W}$ 。Transformer 编码器需要的输入是一个序列, 所以本研究将  $z_0$  的空间维度从多维转化为一维, 从而生成一个新的特征图  $d \times HW$  维。每一个编码器层均有标准架构, 由多头自注意力机制和前馈网络组成。因为 Transformer 的结构是不变的, 所以本研究使用固定位置编码对特征图进行位置信息补充<sup>[18]</sup>。

解码器每一层也使用标准的架构。与 Vaswani 等<sup>[19]</sup>一次预测一个输出不同, 因为我们需要的是空间性的关联而不是时间上的关联, 所以本模型在每个解码器中并行地输出。由于编码器的输出是不变的, 所以  $N$  维特征查询嵌入块必须不同才能产生不同的结果, 这些嵌入块被称为特征查询序列, 和编码阶段类似, 本研究将特征查询序列作为每个解码器的输入。  $N$  维特征查询序列首先经过一个多头自注意力层, 然后和编码器的输出相结合, 共同输入前向传播网络。经过解码输出特征图, 将输出的特征图由一维转化为多维, 再经过  $1 \times 1$  的卷积块将特征图由

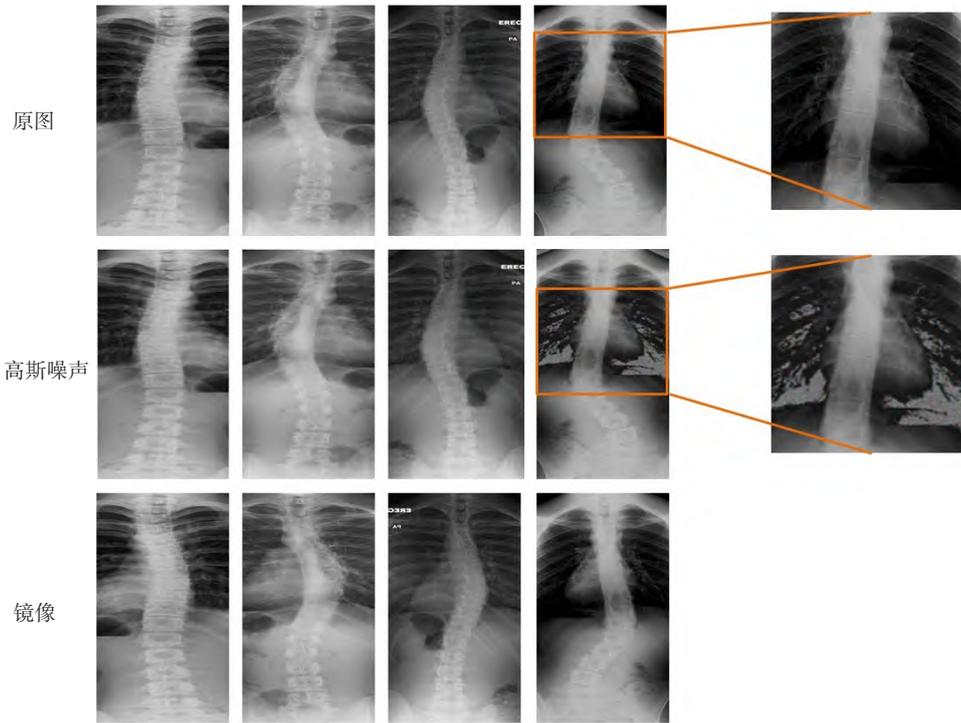


图3 数据增强前后对比

Figure 3 Comparison before and after data augmentation

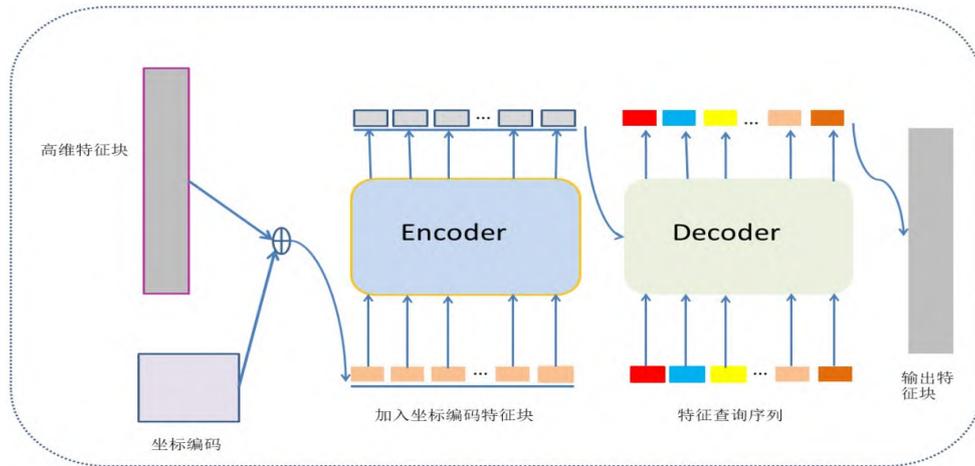


图4 Transformer 模块

Figure 4 Transformer module

$z_0 \in \mathbb{R}^{d \times H \times W}$  转换成  $z_0 \in \mathbb{R}^{C \times H \times W}$ , 最后将整个 transformer 提取的带有空间相关性的特征图经过变形后输出到网络的解码器阶段。

### 1.3 Vector Loss 模块

基于关键点的 Heatmap 常用于姿态估计。对于每一个关键点  $k$ , 它的真实值是一个二维的高斯分布图, 可以用  $e^{-\frac{x^2+y^2}{2\sigma^2}}$  计算。其中半径  $\sigma$  是由每个椎骨大小的尺寸计算得出。本研究使用标准的 Focal Loss 来优化这些参数, Focal Loss 具体形式如下<sup>[16]</sup>:

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

其中,  $i$  代表每一块椎骨的编号,  $N$  代表总共的位置数量, 其中的  $\alpha$  和  $\beta$  使用和文献[16]中相同的 2、4。

当椎骨弯曲特别严重或者图像模糊性比较大时, 模型的预测会出现偏离较大(图 5a)、连续序列中间空缺(图 5b)的现象。为了解决这个问题, 本研究提出 Vector Loss 损失模块(图 5c)。当在预测椎骨的中心点 B 的时候, 模型给出两个候选点 P1 和 P2, 它们与真实标签 B 的 Focal Loss 值相同。通过观察可以看出 P1 与椎骨的曲线趋势更加接近, 应该选择 P1 作为本次的预测, 但是仅使用 Focal Loss 并不能有效地解决这种情况, 模型往往会随机选择一个预测点。

本研究提出的 Vector Loss 损失模块在预测每个椎骨中点时会与相邻的椎骨中心点做一个向量,如图 5c 中的橙色向量所示,将橙色的向量与黑色的真实向量做余弦相似度,便可以很容易去除掉预测点 P2。通过实验,Vector Loss 可以有效地避免图 5a、b 出现的现象。公式如下:

$$L_{\text{vector-loss}} = 1 - \cos(P_{\text{vector}}, GT_{\text{vector}}) \quad (3)$$

其中,  $P_{\text{vector}}$  代表网络的预测向量,  $GT_{\text{vector}}$  代表实际的向量。

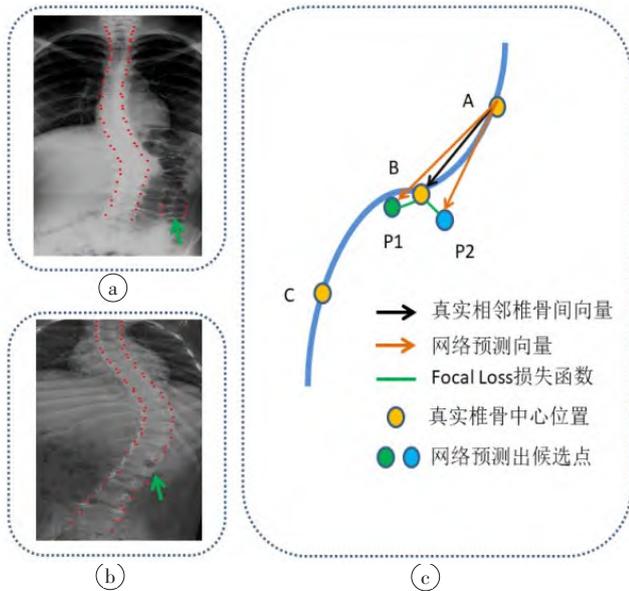


图 5 Vector Loss 模块

Figure 5 Vector Loss module

a: 模型预测时出现错误的偏移预测现象; b: 模型预测时出现连续椎骨遗漏的现象; c: Vector Loss 模块对于异常的预测加以约束

#### 1.4 角偏移量损失函数

在预测阶段,将 Heatmap 放大到原始图像大小时,会因为分辨率的不同而产生浮点数。为了解决浮点数与整数之间的差异,回归一组中心点位移偏差量。真实的中心点与预测偏差量的偏移值为:

$$\varphi = \left( \frac{\vec{c}_x}{s} - \left[ \left| \frac{\vec{c}_x}{s} \right| \right], \frac{\vec{c}_y}{s} - \left[ \left| \frac{\vec{c}_y}{s} \right| \right] \right) \quad (4)$$

其中,  $\vec{c}_x$ 、 $\vec{c}_y$  为网络预测位置;  $s$  为实际位置。

使用 Heatmap 确定椎骨的中心点位置后,使用角偏移量来确定 4 个顶点的位置。角偏移量定义为从椎骨中心指向 4 个顶点的向量。本研究使用  $L_1$  Loss 来优化角偏移量。

## 2 实验

### 2.1 实验平台及数据集

本实验所用的系统是 Ubuntu18.04,使用 PyTorch

编写本研究的代码, RTX 8000 显卡训练本研究的网络。本实验的训练网络框架使用 AdamW 优化器<sup>[20]</sup>, 初始的学习率设置为  $1.5 \times 10^{-4}$ , 基础网络使用预训练的 ResNet, 权重衰减值设置为  $10^{-4}$ , 训练的轮数设置为 1000 轮。

使用 AASCE MICCAI 2019 脊椎挑战赛的公开数据集作为数据集, 包括 609 张训练集和 128 张测试集。使用训练集中的 10% 做验证集, 剩下的 90% 做训练集。本实验使用的数据集是临床经验丰富的骨科医生筛选出各种具有代表性的 X 光图像, 代表了日常临床中各种各样的脊柱疾病案例。

### 2.2 评估指标

使用和 AASCE MICCAI 2019 脊椎挑战赛相同的评价指标。使用平均对称百分比(SMAPE)来评价计算的 Cobb 角:

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

其中,  $i$  代表 PT、MT 及 TL;  $j$  代表第  $j$  张图像;  $N$  代表所有的测试图像数量之和;  $p$  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 来评价预测的好坏, 但本研究还对比评价了 PT、MT、TL 的 SMAPE。

### 2.3 对比实验

将改进的算法与 AASCE MICCAI 2019 脊椎挑战赛排名靠前的论文作对比, 同时比较的对象还包括基于回归算法的方法<sup>[21]</sup>以及基于分割算法的方法<sup>[22]</sup>, 还有同样使用特征点检测算法的方法<sup>[23]</sup>。对于所对比的实验, 使用相同的数据集以及相同的训练轮数。

实验结果如图 6 所示, 其中基础网络和本研究方法包含的两列分别是神经网络输出的向量图和最终的椎骨角点图。基于回归的方法, 因为网络存在大量的学习参数, 在小规模数据集上难以完全学习, 因此预测的 Cobb 角存在较大的偏差; 基于分割的方法在上胸部和主胸部表现较好, 但是在器官堆叠较多的腰部预测上出现了很多错误的预测。基础网络在第 1 行和第 2 行的连续椎骨点预测上出现了错误的预测, 不能准确地预测中间缺失的椎骨; 遇到模糊的图像(第 3 行), 基础网络不能准确地预测出椎骨的位置, 但本研究方法可以准确地定位椎骨的坐标。

本研究所提出的网络对比基于分割的方法具有的优势不仅在于可以预测清晰图像的椎骨角点, 在模糊的 X 光图像上也能准确地预测出脊椎骨的角点坐标; 与基于回归的方法相比, 本研究方法可以在较小的数据集上进行学习训练, 获得更好的效果; 与基础网络相比, 本研究方法可以较好地预测椎骨的相

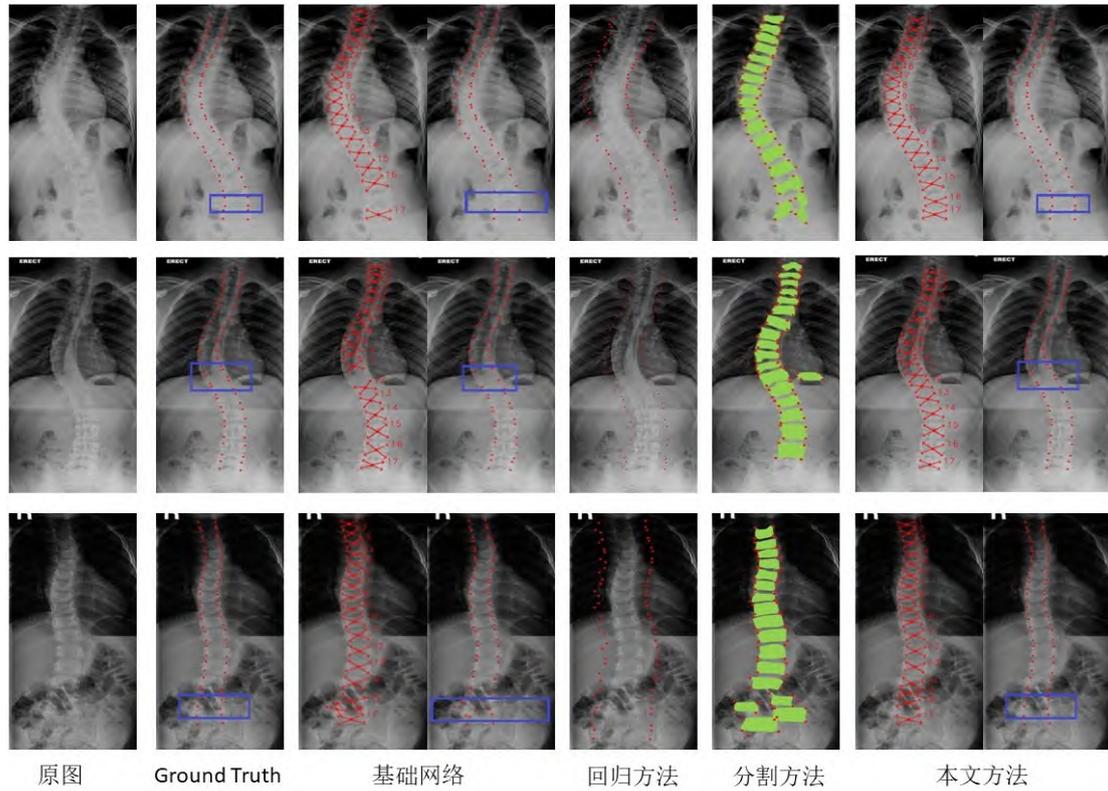


图6 结果比较

Figure 6 Result comparison

邻之间的位置关系,准确地预测脊椎骨的顶点坐标。相对于其他的方法,本文方法具有一定的优势。如表1所示,与基础网络<sup>[23]</sup>相比,本文方法的总体 SMAPE 提高了 1.80, SMAPE<sub>PT</sub>、SMAPE<sub>MT</sub> 和 SMAPE<sub>TL</sub> 分别提升了 0.70、3.86、0.97。分割方法<sup>[22]</sup>在 SMAPE<sub>PT</sub>、SMAPE 优于回归方法<sup>[21]</sup>,即分割方法比回归方法更准确,结果与图6一致。比较回归和分割方法,本文方法具有巨大的优势。由于 AASCE MICCAI 2019 Spine Challenge 仅使用总体 SMAPE 作为指标,本研究将总体 SMAPE 与 Seg4Reg<sup>[24]</sup>和 ACAD<sup>[25]</sup>进行比较,结果也显示本文方法有着更优的性能。

### 2.4 消融实验

为了验证本研究所提出的模型以及功能,采用消融研究的方法去验证各个模块在框架中所起到的作用,消融研究是在基础网络框架之上逐步添加各个模块,实验结果如图7所示。在清晰X光图像上如案例1,所有的模型都可以得到准确的脊椎骨角点预测。在案例2和案例3中,原始的基础网络不能准确地预测椎骨之间两两相互关系,出现了错误的椎骨预测,而本研究提出的在基础网络上嵌入Transformer机制的神经网络框架可以准确地预测椎骨角点的位置。加入Transformer机制对于网络是有很大的提升的,使用Transformer机制可以让神经网络框架学习到各个椎骨之间的相互关系,不会出现连续两块椎骨之间漏掉一块的现象。案例4可以看出对于对比度不高的近腰部位,使用本研究提出的Vector Loss模块可以进行相应的纠正,本研究提出的网络框架可以很好地实现脊椎骨角点的定位。

在消融实验中,本研究使用依次添加模块的方法,在基础网络之上,依次添加Transformer模块、Vector-Loss模块。每个模块对于效果都有明显的提升,尤其是使用Transformer模块之后 SMAPE 提升了 1.19,最后使用的Vector Loss也使得 SMAPE 有 0.61 的提升(表2)。

表1 不同方法的 SMAPE

Table 1 SMAPE of different methods

方法	SMAPE	SMAPE <sub>PT</sub>	SMAPE <sub>MT</sub>	SMAPE <sub>TL</sub>
基础网络 <sup>[23]</sup>	10.81	6.26	18.04	23.42
回归方法 <sup>[21]</sup>	23.43	14.01	21.99	33.01
分割方法 <sup>[22]</sup>	16.48	9.71	25.99	36.01
Seg4Reg <sup>[24]</sup>	22.25	-	-	-
ACAD <sup>[25]</sup>	25.69	-	-	-
本文方法	9.01	5.56	14.18	22.45

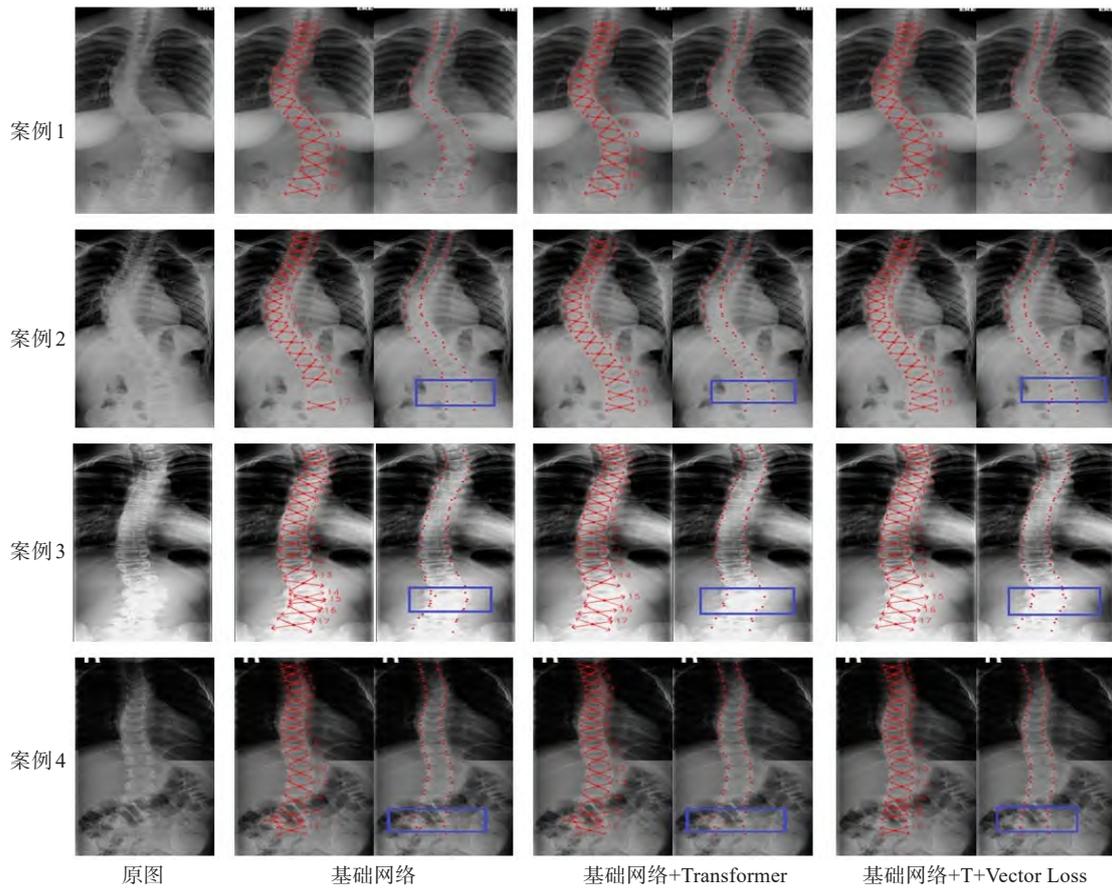


图7 消融实验

Figure 7 Ablation experiment

表2 不同网络方法预测对比

Table 2 Prediction comparison of different network methods

方法	SMAPE
基础网络	10.81
基础网络+Transformer	9.62
基础网络+Transformer + Vector Loss	9.01

### 3 结语

通过上述实验数据对比,本研究提出的嵌入Transformer神经网络框架在脊椎骨定位领域取得了较高的定位精度,可完整地预测脊椎骨的排列走向,进一步预测每一块椎骨的顶点坐标位置,为计算Cobb角奠定坚实基础。本研究方法较好地利用了X光数据集,很好地解决了医学数据集过少的问题;引用Transformer模块很好地解决了原网络只能单独学习每一块椎骨特征,不能学习到脊椎骨之间的相互关系的问题;采用Vector Loss模块很好地纠正了特别模糊时预测偏离的问题。但是添加的Transformer模块参数和运算量略高于基础网络,下一步将会在此基础上深入研究,在不减少预测精度的基础上尽可能减小模型的深度和复杂度。

### 【参考文献】

- [1] Carr AJ. Adolescent idiopathic scoliosis in identical twins[J]. J Bone Joint Surg, 1990, 72(6): 1077.
- [2] Asher MA, Burton DC. Adolescent idiopathic scoliosis: natural history and long term treatment effects[J]. Scoliosis, 2006, 1(1): 1-10.
- [3] Prabhu GK. Automatic quantification of spinal curvature in scoliotic radiograph using image processing[J]. J Med Syst, 2012, 36(3): 1943-1951.
- [4] 任晓静, 颜丽笙. 3种方法测量脊柱侧弯Cobb角的临床应用比较[J]. 福建医药杂志, 2013, 35(6): 121-123.  
Ren XJ, Yan LS. Comparison of 3 methods for measuring Cobb angle of scoliosis[J]. Fujian Medical Journal, 2013, 35(6): 121-123.
- [5] Vrtovec T, Pernuš F, Likar B. A review of methods for quantitative evaluation of spinal curvature[J]. Eur Spine J, 2009, 18(5): 593-607.
- [6] 王伟, 蔡裕兴, 陈卫国, 等. 数字化全脊柱拼接结合核磁共振成像在青少年脊柱侧弯诊治的应用[J]. 中国医学物理学杂志, 2017, 34(6): 594-597.  
Wang W, Cai YX, Chen WG, et al. Value of digital whole spine splicing combined with magnetic resonance imaging in diagnosis and treatment of adolescent scoliosis[J]. Chinese Journal of Medical Physics, 2017, 34(6): 594-597.
- [7] 姚红艳, 邓兴, 陈晓飞, 等. 腰椎X线摄影人工智能测量技术研究进展[J]. 中国医学物理学杂志, 2021, 38(12): 1579-1584.  
Yao HY, Deng X, Chen XF, et al. Advances in artificial intelligence technology for parameter measurement in lumbar X-ray photograph[J]. Chinese Journal of Medical Physics, 2021, 38(12): 1579-1584.
- [8] Wills BP, Auerbach JD, Zhu X, 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.
- [9] Prabhu GK. Automatic quantification of spinal curvature in scoliotic radiograph using image processing[J]. J Med Syst, 2012, 36(3): 1943-1951.

- [10] Anitha H, Karunakar AK, Dinesh KV. Automatic extraction of vertebral endplates from scoliotic radiographs using customized filter [J]. *Biomed Eng Lett*, 2014, 4(2): 158-165.
- [11] Zhen X, Wang Z, Islam A, et al. Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation [J]. *Med Image Anal*, 2016, 30: 120-129.
- [12] Sun HL, Zhen X, Bailey C, et al. Direct estimation of spinal Cobb angles by structured multi-output regression [C]//International Conference on Information Processing In Medical Imaging. Cham: Springer, 2017: 529-540.
- [13] Wu H, Bailey C, Rasoulinejad P, et al. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2017: 127-135.
- [14] Wang S, Huang S, Wang L. Spinal curve guide network (SCG-Net) for accurate automated spinal curvature estimation [C]//International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging. Cham: Springer, 2019: 107-112.
- [15] Ronneberger O, Fischer P, Brox T, et al. U-net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer, 2015: 234-241.
- [16] Abolfazl A, Pradhan B, Alamri A. VNet: an end-to-end fully convolutional neural network for road extraction from high-resolution remote sensing data [J]. *IEEE Access*, 2020, 8: 179424-179436.
- [17] Yi J, Wu P, Huang Q, et al. Multi-scale cell instance segmentation with keypoint graph based bounding boxes [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 369-377.
- [18] Bello I, Zoph B, Vaswani A, et al. Attention augmented convolutional networks [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. South Korea: IEEE, 2019: 3286-3295.
- [19] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]. *Advances in Neural Information Processing Systems*, 2017: 5998-6008.
- [20] Loshchilov I, Hutter F. Decoupled weight decay regularization [J]. *arXiv preprint*, 2017. arXiv: 1711.05101.
- [21] Wang L, Xu Q, Leung S, et al. Accurate automated Cobb angles estimation using multi-view extrapolation net [J]. *Med Image Anal*, 2019, 58: 101542.
- [22] Horng MH, Kuok CP, Fu MJ, et al. Cobb angle measurement of spine from X-ray images using convolutional neural network [J]. *Comput Math Method Med*, 2019. <https://doi.org/10.1155/2019/6357171>.
- [23] Yi J, Wu P, Huang Q, et al. Vertebra-focused landmark detection for scoliosis assessment [C]//2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, 2020: 736-740.
- [24] Lin Y, Zhou HY, Ma K, et al. Seg4Reg networks for automated spinal curvature estimation [C]//International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging. Cham: Springer, 2019: 69-74.
- [25] Khanal B, Dahal L, Adhikari P, et al. Automatic Cobb angle detection using vertebra detector and vertebra corners regression [C]//International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging. Cham: Springer, 2019: 81-87.

(编辑:谭斯允)