

# 基于SE-SWTNet的孕妇引产超声图像的关键点自动检测研究

陈黄菲<sup>1\*</sup>, 陈 胜<sup>1#</sup>, 姚丽萍<sup>2</sup>

<sup>1</sup>上海理工大学光电信息与计算机工程学院, 上海

<sup>2</sup>上海市第一妇婴保健院超声科, 上海

收稿日期: 2025年3月1日; 录用日期: 2025年3月25日; 发布日期: 2025年3月31日

## 摘 要

为了降低引产风险并提高妊娠临床诊断的准确性, 需要识别产前超声检查中的几个关键参数。收集妊娠第38周至42周的120幅宫颈超声图像, 对数据进行标注和预处理。以传统swin transformer (Shifted Window Transformer)网络为基线网络, 引入转置卷积和最后的输出特征层, 能直接输出关键点参数; 同时加入正则化函数和SE (Squeeze and Excitation)注意力机制, 减少过拟合风险, 提高模型的泛化能力。将SE-SWTNet (Squeeze and Excitation Shifted Window Transformer)与现有的4种算法进行对比。结果显示, SE-SWTNet能较为精确地捕捉宫颈超声图像中的关键点参数。在测试集上, SE-SWTNet的平均半径误差为1.11 mm, 并分别在4 mm、5 mm和6 mm误差范围内达到了93.5%、96.5%和97.5%的准确率。SE-SWTNet可以实现基于孕妇引产超声图像的高精度关键点自动检测的功能, 为临床医生提供更准确的诊断信息, 从而降低风险, 提高成功率。

## 关键词

引产, 关键点检测, 宫颈超声影像, 注意力机制

# Automated Key Point Detection for Labor Induction Ultrasound Imaging in Pregnant Women Using SE-SWTNet

Huangfei Chen<sup>1\*</sup>, Sheng Chen<sup>1#</sup>, Liping Yao<sup>2</sup>

<sup>1</sup>School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

<sup>2</sup>Ultrasound Department, Shanghai First Maternity and Infant Hospital, Shanghai

\*第一作者。

#通讯作者。

文章引用: 陈黄菲, 陈胜, 姚丽萍. 基于 SE-SWTNet 的孕妇引产超声图像的关键点自动检测研究[J]. 建模与仿真, 2025, 14(4): 9-18. DOI: 10.12677/mos.2025.144260

## Abstract

In order to reduce the risk of induced abortion and improve the accuracy of clinical diagnosis of pregnancy, it is necessary to identify several key parameters in prenatal ultrasound examination. Collect 120 cervical ultrasound images from weeks 38 to 42 of pregnancy, annotate and preprocess the data. Using the traditional Swin transformer (Shifted Window Transformer) network as the baseline network, introducing transposed convolution and the final output feature layer, it can directly output keypoint parameters; Simultaneously incorporating regularization functions and SE (Squeeze and Excitation) attention mechanisms to reduce the risk of overfitting and improve the model's generalization ability. Compare SE SWTNet (Squeeze and Excitation Shifted Window Transformer) with four existing algorithms. The results showed that SE SWTNet can accurately capture key point parameters in cervical ultrasound images. On the test set, the average radius error of SE SWTNet is 1.11 mm, and it achieved accuracies of 93.5%, 96.5%, and 97.5% within error ranges of 4 mm, 5 mm, and 6 mm, respectively. SE SWTNet can achieve high-precision automatic detection of key points based on ultrasound images of induced abortion in pregnant women, providing more accurate diagnostic information for clinical doctors, thereby reducing risks and improving success rates.

## Keywords

Induction of Labor, Key Point Detection, Cervical Ultrasound Images, Attention Mechanism

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在产科，引产(Induction of Labour, IOL)是妊娠晚期孕妇在自然分娩开始前做采取的一种人工干预措施[1]，是终止妊娠并实现阴道分娩的一种常见方法。目前在美国，大约 25%的孕晚期妊娠孕妇需要进行引产[2]。引产的主要目的是终止可能对母体或胎儿造成损害的病理情况，例如胎膜早破、先兆子痫、胎儿宫内生长受限、羊水异常，妊娠期糖尿病[3]等。然而，引产后也会出现产程延长甚至无效产程[4]的现象，最终导致剖宫产风险增加。因此，为了提高引产成功率并降低剖宫产率，预测引产结局，寻找选择合适且安全的引产方法成为了产科研究热点。

在影像技术普及之前，宫颈成熟度的评估主要依靠人工检查，并通过 Bishop 评分[5]来进行评定。然而，这种方法存在主观性强以及观察者之间和观察者自身评价不一致的局限性。因此，需要一种更为客观的评估方法。现如今，超声影像已成为孕妇产前检查中不可或缺的工具，经阴道超声(TVS) [6]作为一种客观的宫颈评估方法[7]，对评估早产风险，预测引产能否成功乃是引产后的妊娠结局[8]起到了重要的作用。本研究旨在识别 TVS 评分中宫颈长度(Cervical length, CL) [9]、前宫颈前角(Anterior lip of cervix, ACA) [10]和后宫颈角(Posterior fornix of cervix, PCA) [11]这三个参数，以帮助产科医生评估宫颈的成熟度，从而预测孕妇是否需要引产。

计算机技术的发展，深度学习技术使得上述需求成为可能。通过模型自动学习超声影像中的特征，识别并自动标识医生所需的信息，极大地提高了诊断的准确性和效率。本研究提出了一种宫颈超声图像

关键点自动检测算法, 能够实现自动、定量和精确地定位到所需参数的关键点。将标注好的图像数据输入到本研究构建的深度学习模型中进行训练, 通过模型训练得到关键点的训练标签, 随机抽取未标注的图像数据进行预测, 利用训练好的模型自动预测出图像中的关键点。最后, 我们将预测的关键点与医生手动标注的关键点进行比较和验证。在模型构建过程中, 选取性能较好的 swin transformer 为基线网络, 引入反卷积层作为上采样层以提高图像的空间分辨率。同时, 为了增强模型的泛化能力和抑制过拟合现象, 加入 Dropout 函数和注意力机制。构建了一种关键点检测网络 SE-SWTNet, 该网络能自动识别宫颈超声图像中的特征, 并标记影响引产结局的几个参数的关键点。

## 2. 资料与方法

### 2.1. 数据集

本研究所用的实验数据均来源于上海市第一妇婴保健院, 使用二维经会阴超声测量妊娠第 38 周至 42 周的妇女, 获取宫颈超声图像, 数据收集的时间为 2022 年 11 月至 2023 年 7 月。剔除质量较差, 特征不明显的图像后, 共采集了 120 张宫颈超声图像。本研究经医院伦理委员会批准[编号(No): KS2479], 患者均知情同意。

### 2.2. 数据预处理

对 123 名受试者的宫颈超声图像进行评估, 每张图像由临床专家标注。对于实验中要用到的标签数据, 利用 Labelme 工具标注后经医生审核并通过。其中 70% 的数据集被随机选择用于训练, 剩下的 30% 用于测试。对于原始图像数据, 所有图像体积都处理为各向同性, 并调整为  $448 \times 448$  像素, 垂直和水平方向上的像素间距为 0.146 mm。模型预测的关键点位置将以坐标形式存储。两个关键点组成宫颈管长度 (Cervical Length, CL), 三个关键点组成宫颈前角 (Anterior Lip of Cervix, ACA), 另三个关键点组成宫颈后角 (Posterior Fornix of Cervix, PCA)。因此, 这三个用于判断引产结局的参数由 4 个关键点来确定。见图 1。

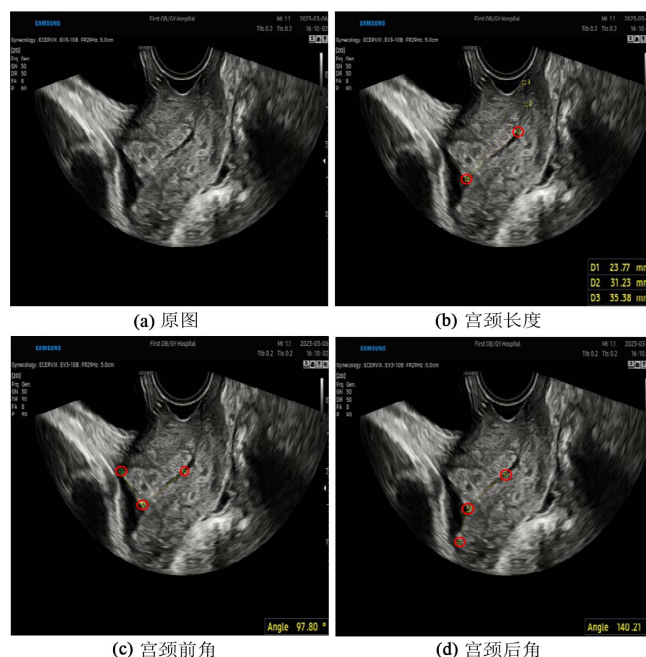


Figure 1. Training sample display

图 1. 训练样本展示

### 3. 网络结构

SE-SWTNet 的网络结构包括编码器和解码器两部分，编码器由 swin transformer 的前四个阶段构成，用于从输入的图像中提取特征并进行全局特征融合。解码器部分负责将编码器输出的特征图转换为最终的输出，添加了转置卷积(Transposed Convolution)层，通过对特征图进行填充和卷积操作，扩大特征图的尺寸，恢复原始图像的细节。同时，还添加了注意力模块，增强模型的表征能力，使其更好地捕获图像中的关键信息。为了使输出图像的通道数达到与原始标注图像的关键点个数保持一致，在网络的最后添加了一个卷积层。最终的卷积层将经过残差块和转置卷积后的特征映射转换为与关键点数量相匹配的输出特征映射，其中每个通道对应一个关键点的预测。整个网络是端到端的，可以直接接收原始图像作为输入，并输出直观的预测结果。见图 2。

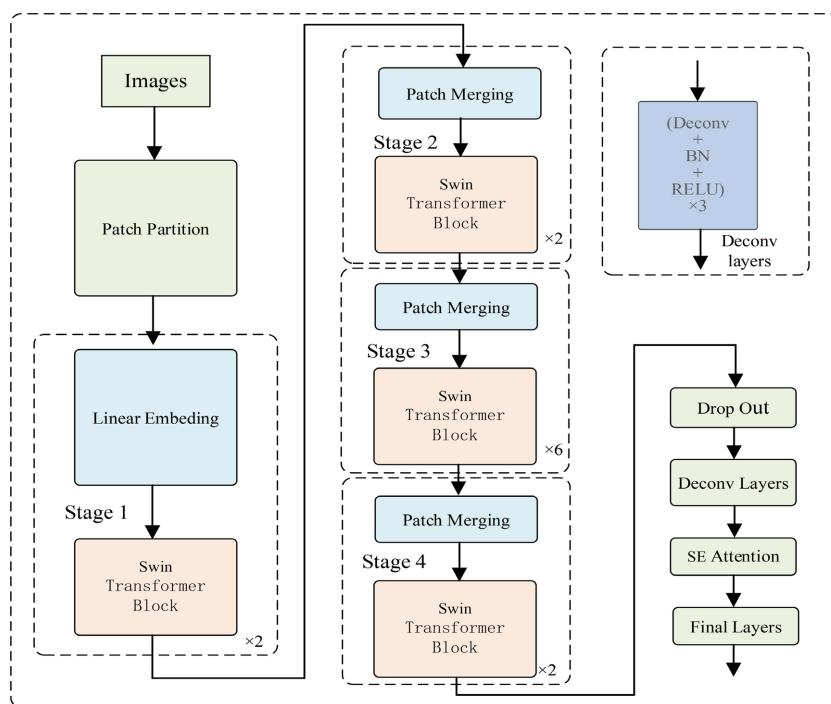


Figure 2. Network structure diagram  
图 2. 网络结构图

#### 3.1. Swin Transformer

Swin Transformer [12]是一种自注意力机制的创新变体，网络结构如图 3 所示。旨在解决传统 Vision Transformer (ViT) [13]在处理高分辨率图像时面临的计算复杂度问题。与传统 ViT 不同，Swin Transformer 通过引入窗口自注意力机制，限制了每个自注意力计算的范围，从而大大减少了计算和内存的消耗。

为了实现跨窗口的信息交互，Swin Transformer 采用了位移窗口(Shifted Window)。这种设计通过在每一层的窗口内进行局部自注意力计算，并在下一层中通过平移窗口来连接不同窗口之间的信息。这种机制使得网络能够在保持计算效率的同时捕捉到全局特征，避免了传统自注意力机制在计算上的高消耗。

Swin Transformer 的基本结构由多个 Swin Block 组成，每个 Swin Block 包含两个核心模块：窗口化自注意力(Window-Based Self-Attention)和多层感知机。通过这些 Swin Block 的堆叠，网络能够逐层提取图像的多层次特征。此外，Swin Transformer 采用了分层设计的架构，将图像从高分辨率的块逐渐转化为

低分辨率的块,从而能够学习从局部到全局的特征表示。这种分层策略使得 Swin Transformer 不仅在计算效率上具有优势,而且能够适应不同尺寸和分辨率的图像输入。

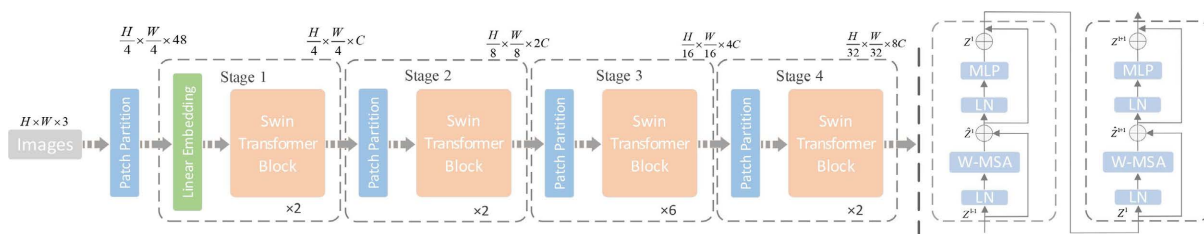


Figure 3. Swin transformer network structure diagram

图 3. Swin transformer 网络结构图

### 3.2. 转置卷积

为使模型的输出尺寸与输入的图像保持一致,在模型的解码器部分加入转置卷积。转置卷积可以通过编码器输出的小尺寸图像恢复到卷积操作前的尺寸。它通过对输入的特征矩阵进行上采样,填充特征图的外层,并进行一系列的卷积操作,将卷积核参数逆向翻转,最终扩大特征矩阵的尺寸,以便进行像素级的预测。本研究设置了三个转置卷积层,每层都包含一个转置卷积、批归一层和 ReLU 激活函数,这样可以增加模型的非线性和稳定性。见图 4。

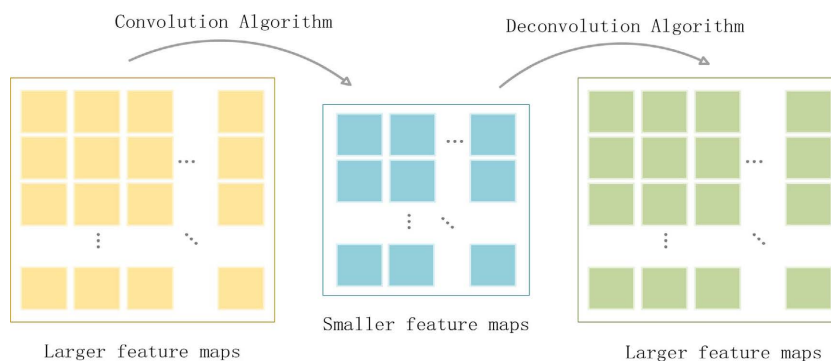


Figure 4. Transposed convolution

图 4. 转置卷积

### 3.3. SE 注意力机制

在转置卷积之后,为提高网络的表征能力,使模型能够关注信息量最大的通道特征并抑制不重要的特征,本研究引入了 SE 注意力机制[14],该机制的主要结构包括三个部分。第一部分,输入的特征矩阵经过一个全局池化层,压缩其二维特征,输出大小为  $1 \times 1 \times 256$  的特征矩阵;第二部分由全连接层、激活函数和全连接层复合构成,将通道数先降维后激活,最后恢复维度,特征矩阵由  $1 \times 1 \times 256$  变为  $1 \times 1 \times 64$  最后恢复到  $1 \times 1 \times 256$ ,以降低模型复杂度;第三部分,在完成两次全连接之后,使用 sigmoid 将值固定在  $[0, 1]$  之间,以获取每一个通道的权值,依次将权值系数和特征矩阵相乘,最终恢复输入大小的特征矩阵。见图 5。

### 3.4. 模型后处理

现有方法常应用高斯热图(Gaussian Heatmap) [15]来表示每个关键点的概率分布,高斯热图是一种

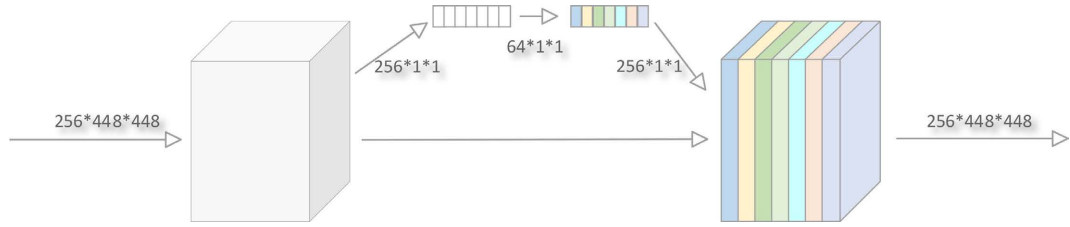


Figure 5. SE attention mechanism  
图 5. SE 注意力机制

平滑的概率分布，其具有连续性和平滑性，通过求最大值所在的位置坐标，可以更好地表示关键点的位置信息。为从预测的热图中提取 2D 关键点坐标，使用可微分的 `soft-argmax` [16] 来减少由于热图量变化误差带来的误差。具体而言，将关键点的热图作归一化，使下一步的加权求和范围在 0~1 之间，见公式(1)。

$$\Phi(h_{i,j}) = \frac{e^{h_{i,j}}}{\sum_{K=1}^W \sum_{l=1}^H e^{h_{k,l}}} \quad (1)$$

式中， $h_{i,j}$  是关键点位置  $(i,j)$  处的热图的值， $W \times H$  是热图的大小。分别用  $x, y$  两个与热图大小一样的卷积核对归一化后的热图加权求和，等价于对应元素相乘再相加，等价于卷积运算，见公式(2)。

$$\Psi_d(h) = \sum_{i=1}^W \sum_{j=1}^H W_{i,j,d} \Phi(h_{i,j}) \quad (2)$$

式中，其中  $d$  是一个给定的分量  $x$  或  $y$ ， $W$  是一个对应于坐标  $(x,y)$  的  $W \times H \times 2$  权值矩阵。矩阵  $W$  可以用其分量  $W_x$  和  $W_y$  表示，它们是二维离散归一化斜坡，见公式(3)。

$$W_{i,j,x} = \frac{i}{W}, W_{i,j,y} = \frac{j}{H} \quad (3)$$

最后，给定一个热图  $h$ ，给出了预测关键点的位置为公式(4)。

$$y = (\Psi_x(h), \Psi_y(h))^T \quad (4)$$

本研究所用的损失函数是 pixel-wise 的 Mean Squared Error (MSE)，见公式(5)。

$$L_{MSE} = \frac{\sum_{i=1}^K \sum_{y=1}^W \sum_{x=1}^H [h_i(x,y) - \hat{h}_i(x,y)]^2}{K} \Bigg/ \frac{W}{H} \quad (5)$$

式中  $K$  为关键点的个数， $h_i$  和  $\hat{h}_i$  分别代表对应的真实点和预测点的值。

### 3.5. 网络训练

实验环境为：Intel Core i9-10940X，NVIDIA GTX 3090 显卡，64 位操作系统，Pytorch 深度学习框架。编程平台为 Pycharm，图形加速环境为 CUDA11.1。实验设置训练轮数为 200 个 epoch，初始学习率为 0.001。采用 Adam 优化器根据历史梯度信息来自适应地调整学习率，批次大小为 8。为了加速模型收敛，选取轻量级的 swin\_tiny 预训练权重。

### 3.6. 评价指标

对于预测得到的关键点，采用半径误差(RE)，平均半径误差(MRE)，标准差(SD)对模型的性能进行量

化评估。其中，半径误差采用欧几里得距离进行测量，见公式(5)~(9)。

$$R_i = \left| (x_m, y_m) - (x_n, y_n) \right|_2 \quad (6)$$

$$S_{\text{MRE}} = \frac{\sum_{i=1}^{\delta} R_i}{\delta} \quad (7)$$

$$T_{\text{SD}} = \sqrt{\frac{\sum_{i=1}^{\delta} (R_i - S_{\text{MRE}})^2}{\delta - 1}} \quad (8)$$

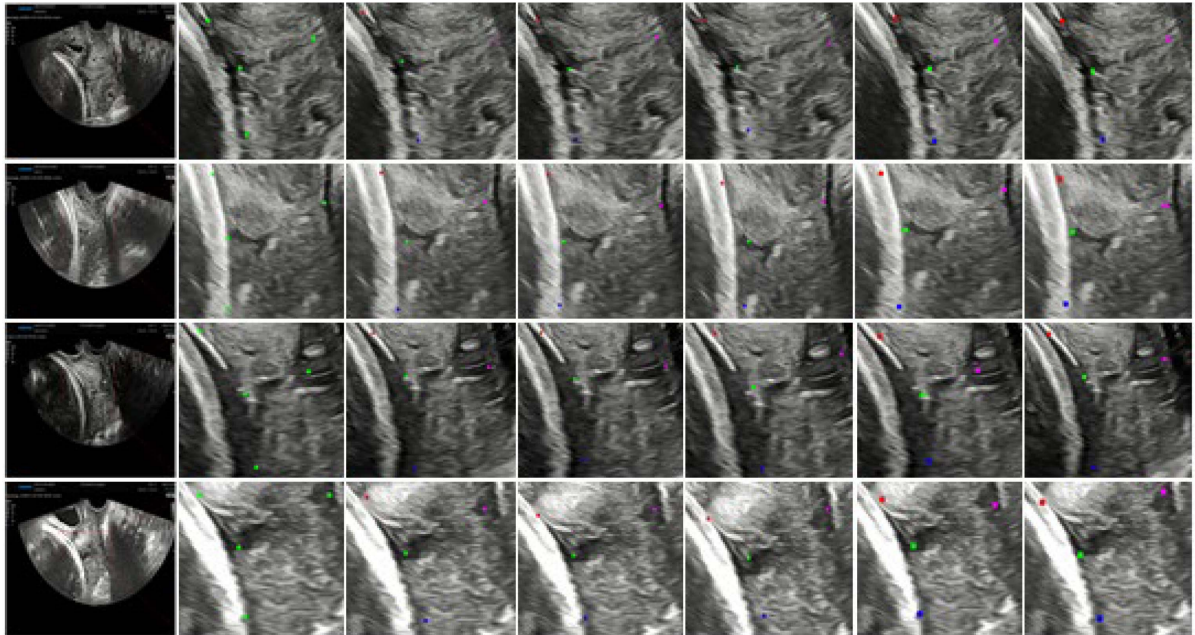
$$T_{\omega} = \frac{\#\left(\left\{y_t : R_i \leq \omega\right\}\right)}{\delta} \quad (9)$$

其中  $R_i$  表示半径误差， $(x_m, y_m)$  表示预测关键点的坐标， $(x_n, y_n)$  表示医生标注关键点的坐标， $\delta$  代表测试总样本的数量与每个样本中关键点数量的乘积， $T_{\omega}$  表示半径误差小于指定范围内的关键点数量在所有关键点数量中的占比， $\omega$  表示临床上可以接受的误差范围，分别设置 4 mm，5 mm，6 mm 三个门限值对预测关键点的精度进行评判。

## 4. 结果

### 4.1. 主观评价

实验的可视化结果分析如下，选取 4 幅具有不同纹理特征与边缘轮廓的宫颈超声图像与 HRNet、ViTPose、U-Net 和 ResNet 进行对比。从图像上可以清晰地观察到，本研究的方法相较于其他四个方法，所预测的关键点与医生标注的关键点更加接近，表现出更高的精度和准确性。见图 6。



**Figure 6.** Key point prediction results  
**图 6.** 关键点预测结果

4.2. 客观评价

**Table 1.** Comparison of key point prediction results  
**表 1.** 关键点预测结果对比

| 方法      | 平均半径误差/mm | 标准差/mm | 不同误差范围内的准确率/% |      |      |
|---------|-----------|--------|---------------|------|------|
|         |           |        | 4 mm          | 5 mm | 6 mm |
| U-Net   | 3.31      | 2.10   | 63.3          | 73.6 | 82   |
| ResNet  | 3.32      | 1.97   | 71.7          | 83.2 | 87.5 |
| HRNet   | 2.27      | 1.21   | 72.8          | 83.8 | 88   |
| VITPose | 1.62      | 1.02   | 80.3          | 83   | 92.8 |
| Ours    | 1.11      | 0.70   | 93.5          | 96.5 | 97.5 |

由表 1 可得, SE-SWTNet 在平均半径误差、标准差以及半径误差小于指定范围内的关键点数量占比等指标上均优于其他三个模型。设置的三个门限范围, U-Net 的关键点识别准确率最低, 原因可能是 U-Net 的模型结构过于简单, 对于处理复杂图像的关键点识别任务没有充分提取图像的特征信息, 导致解码器部分无法准确恢复关键点的位置信息。VITPose 虽然获得了仅次于本文方法的准确率, 但相较于传统的 CNN 架构, 其利用 Transformer 作为主干网络, 在处理医学图像关键点检测任务中具有优异的全局感知能力。这使得 VITPose 能够更好地捕捉到图像数据中不同关系的依赖, 有助于更好地理解医学图像复杂的结构, 为后续研究医学图像关键点检测任务提供了有力的理论依据。HRNet 的准确率位于 U-Net 和 VITPose 之间, 它主要通过构建不同分辨率的子网络来融合多维度信息, 然而其精度可能受到网络结构较浅的限制。尽管采用了密集连接的设计, 但在处理复杂的医学图像时, 可能需要更深层次的网络结构来提高特征表征能力。如果网络深度不足, 可能会导致模型无法有效地学习到医学图像中的抽象特征, 从而影响关键点检测的准确性。

为评估 SE-SWTNet 在实际临床环境中的性能, 我们在测试集中引入了不同水平的高斯噪声( $\sigma = 0.01, 0.05, 0.1$ )以及高斯模糊(kernel size = 3, 5, 7), 并测量关键点检测的精度变化。实验结果表明, SE-SWTNet 在噪声和伪影干扰下仍能保持较高的检测精度, 见表 2。这表明该模型对噪声和伪影具有较强的鲁棒性, 能够有效应对实际临床场景中的多种干扰因素。

**Table 2.** Comparison of key point detection performance under different noise and fuzzy interference  
**表 2.** 不同噪声和模糊干扰下的关键点检测性能对比

| 干扰类型 | 参数值             | 平均半径误差 $\pm$<br>标准差(mm) | 不同误差范围内的准确率/% |      |      |
|------|-----------------|-------------------------|---------------|------|------|
|      |                 |                         | 4 mm          | 5 mm | 6 mm |
| 无干扰  | -               | 1.11 $\pm$ 0.70         | 93.5          | 96.5 | 97.5 |
|      | $\sigma = 0.01$ | 1.15 $\pm$ 0.74         | 92.8          | 95.9 | 97.2 |
|      | $\sigma = 0.05$ | 1.32 $\pm$ 0.89         | 89.5          | 93.2 | 95.6 |
|      | $\sigma = 0.1$  | 1.58 $\pm$ 1.05         | 85.6          | 90.1 | 94.3 |
| 高斯噪声 | kernel size = 3 | 1.18 $\pm$ 0.76         | 92.3          | 95.5 | 97.0 |
|      | kernel size = 5 | 1.35 $\pm$ 0.91         | 88.7          | 92.8 | 95.2 |
|      | kernel size = 7 | 1.61 $\pm$ 1.08         | 84.9          | 89.7 | 93.8 |

### 4.3. 消融实验

为探究实验加入的各个模块对关键点识别结果的影响,设计了消融实验,将 Dropout、MSE 损失函数、SE 注意力机制加入到基础网络 swin transformer 中。可以看到,单独加入 Dropout 或 SE 注意力机制均能有效提高关键点识别的准确率和降低平均半径误差。当同时加入 Dropout、损失函数和 SE 注意力机制时,模型的表现进一步提升,取得了最优的结果:平均半径误差降低至 1.11 mm,且在 4 mm、5 mm 和 6 mm 误差范围内的准确率分别达到了 93.5%、96.5%和 97.5%。见表 3。

**Table 3.** Ablation experiment

**表 3.** 消融实验

| 方法                                    | 平均半径误差 $\pm$<br>标准差(mm) | 不同误差范围内的准确率/% |      |      |
|---------------------------------------|-------------------------|---------------|------|------|
|                                       |                         | 4 mm          | 5 mm | 6 mm |
| SWT                                   | 1.45 $\pm$ 1.19         | 86.5          | 90.2 | 93.3 |
| SWT + Dropout                         | 1.32 $\pm$ 0.92         | 88            | 91.8 | 94.1 |
| SWT t+ L <sub>MSE</sub>               | 1.28 $\pm$ 0.81         | 89.7          | 93.5 | 95.7 |
| SWT + SE                              | 1.22 $\pm$ 0.78         | 91.2          | 94.8 | 96.3 |
| SWT + Dropout + L <sub>MSE</sub> + SE | 1.11 $\pm$ 0.70         | 93.5          | 96.5 | 97.5 |

## 5. 总结

近年来,关键点检测在医学图像处理中取得了较多的研究进展,在许多其他疾病的诊断过程中,关键点检测技术被广泛应用。如 Ramadan [17]等人通过构建神经网络模型,实现了在侧位头颅测量 X 射线图像上自动检测头颅测量标志的功能。Fard [18]等人提出的编码器-解码器网络 PoseNet,用于自动检测颈椎矢状面标志点。这些研究表明,关键点检测在医学图像分析中的潜力巨大,能够显著提升诊断的效率和准确性,为临床医生提供强有力的辅助支持。

对于孕妇引产前宫颈图像参数的自动识别,由于孕晚期(38~42 周)的宫颈超声图像具有极高的复杂度,因为宫颈在这一阶段会发生显著的变化,包括宫颈长度的缩短、软化和扩张。这些变化不仅难以通过肉眼判断,还因图像中包含的噪声和其他生理结构(如胎头和羊水)的干扰,使得关键点自动检测变得尤为困难。本研究选取 swin transformer 为主干网络,并引入了转置卷积和注意力机制,在一定程度上实现了网络在复杂医学图像上的关键点精准识别。实验结果表明,SE-SWTNet 在关键点识别精度上优于 HRNet、VITPose、U-Net 和 ResNet,能够较为稳定、准确地识别宫颈图像的关键点参数。

然而,本研究的方法仍存在一些局限性。例如,实验采用的数据量较少,更多的数据可能会训练出更为精准的结果;此外,实验所用数据为二维超声数据,这在一定程度上限制了模型的性能。三维超声数据包含了丰富的空间信息,而二维图像则可能丢失了部分关键信息。直接对三维数据进行关键点检测,可以更全面地利用图像中的空间信息,提高检测的准确性和稳定性。

本研究预测的孕妇引产前宫颈图像关键点参数对产科医生的阅片以及判断宫颈成熟度有着很大的帮助,但宫颈管长度、前宫颈前角和后宫颈角这三个参数对预测引产成功与否及引产后的妊娠结局仍需经过长期的时间和实践考量。

## 基金项目

国家自然科学基金 81101116。

## 参考文献

- [1] Yang, Q., Zhou, C., Chen, Y., Pei, J., Hua, X. and Yao, L. (2023) Prediction Model for Successful Induction of Labor by Cervical Strain Elastography Diagnosed at Late-Term Pregnancy in Nulliparous Women: A Prospective Cohort Study. *BMC Pregnancy and Childbirth*, **23**, Article No. 114. <https://doi.org/10.1186/s12884-023-05426-7>
- [2] Walani, S.R. (2020) Global Burden of Preterm Birth. *International Journal of Gynecology & Obstetrics*, **150**, 31-33. <https://doi.org/10.1002/ijgo.13195>
- [3] Rydahl, E., Eriksen, L. and Juhl, M. (2019) Effects of Induction of Labor Prior to Post-Term in Low-Risk Pregnancies: A Systematic Review. *JBIM Database of Systematic Reviews and Implementation Reports*, **17**, 170-208. <https://doi.org/10.11124/jbisrir-2017-003587>
- [4] Troyanova-Wood, M., Meng, Z. and Yakovlev, V.V. (2019) Differentiating Melanoma and Healthy Tissues Based on Elasticity-Specific Brillouin Microspectroscopy. *Biomedical Optics Express*, **10**, 1774-1781. <https://doi.org/10.1364/boe.10.001774>
- [5] Zhou, Y., Jin, N., Chen, Q., Lv, M., Jiang, Y., Chen, Y., et al. (2021) Predictive Value of Cervical Length by Ultrasound and Cervical Strain Elastography in Labor Induction at Term. *Journal of International Medical Research*, **49**, 1-17. <https://doi.org/10.1177/0300060520985338>
- [6] Agrawal, A., Tripathi, P.S., Bhandari, G., Kheti, P., Madhpuriya, G. and Rathore, R. (2022) Comparative Study of TVS Cervical Score and Bishop Score in Prediction of Successful Labour Induction. *Egyptian Journal of Radiology and Nuclear Medicine*, **53**, Article No. 138. <https://doi.org/10.1186/s43055-022-00794-0>
- [7] Garg, P., Gomez Roig, M.D. and Singla, A. (2019) An Ultrasound Prediction Model for Probability of Vaginal Delivery in Induction of Labor. *Perinatal Journal*, **27**, 161-168. <https://doi.org/10.2399/prn.19.0273007>
- [8] Lu, J., Cheng, Y.K.Y., Ho, S.Y.S., Sahota, D.S., Hui, L.L., Poon, L.C., et al. (2019) The Predictive Value of Cervical Shear Wave Elastography in the Outcome of Labor Induction. *Acta Obstetrica et Gynecologica Scandinavica*, **99**, 59-68. <https://doi.org/10.1111/aogs.13706>
- [9] Li, P., Tsui, W.L. and Ding, D. (2023) The Association between Cervical Length and Successful Labor Induction: A Retrospective Cohort Study. *International Journal of Environmental Research and Public Health*, **20**, Article No. 1138. <https://doi.org/10.3390/ijerph20021138>
- [10] Park, H.S., Kwon, H., Kwak, D.W., Kim, M.Y., Seol, H., Hong, J., et al. (2019) Addition of Cervical Elastography May Increase Preterm Delivery Prediction Performance in Pregnant Women with Short Cervix: A Prospective Study. *Journal of Korean Medical Science*, **34**, e68. <https://doi.org/10.3346/jkms.2019.34.e68>
- [11] Kamel, R., Garcia, F.S.M., Poon, L.C. and Youssef, A. (2021) The Usefulness of Ultrasound before Induction of Labor. *American Journal of Obstetrics & Gynecology MFM*, **3**, Article ID: 100423. <https://doi.org/10.1016/j.ajogmf.2021.100423>
- [12] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 10012-10022. <https://doi.org/10.1109/iccv48922.2021.00986>
- [13] Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., et al. (2023) A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 87-110. <https://doi.org/10.1109/tpami.2022.3152247>
- [14] Deng, J., Ma, Y., Li, D., Zhao, J., Liu, Y. and Zhang, H. (2020) Classification of Breast Density Categories Based on Se-Attention Neural Networks. *Computer Methods and Programs in Biomedicine*, **193**, Article ID: 105489. <https://doi.org/10.1016/j.cmpb.2020.105489>
- [15] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., et al. (2021) CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronics*, **10**, Article No. 2470. <https://doi.org/10.3390/electronics10202470>
- [16] Ahn, D., Kim, S., Hong, H. and Chul Ko, B. (2023) Star-Transformer: A Spatio-Temporal Cross Attention Transformer for Human Action Recognition. 2023 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 2-7 January 2023, 3330-3339. <https://doi.org/10.1109/wacv56688.2023.00333>
- [17] Ramadan, R.A., Khedr, A.Y., Yadav, K., Alreshidi, E.J., Sharif, M.H., Azar, A.T., et al. (2022) Convolution Neural Network Based Automatic Localization of Landmarks on Lateral X-Ray Images. *Multimedia Tools and Applications*, **81**, 37403-37415. <https://doi.org/10.1007/s11042-021-11596-3>
- [18] Fard, A.P., Ferrantelli, J., Dupuis, A. and Mahoor, M.H. (2022) Sagittal Cervical Spine Landmark Point Detection in X-Ray Using Deep Convolutional Neural Networks. *IEEE Access*, **10**, 59413-59427. <https://doi.org/10.1109/access.2022.3180028>