

## 引文格式:

黄凌风, 杨世龙, 谢耀钦. YOLO-PointMap: 基于轻量化动态特征融合的实时人体背部穴位识别 [J]. 集成技术, 2025, 14(2): 58-70.

Huang LF, Yang SL, Xie YQ. YOLO-PointMap: real-time human back acupoint recognition based on lightweight dynamic feature fusion [J]. Journal of Integration Technology, 2025, 14(2): 58-70.

## YOLO-PointMap: 基于轻量化动态特征融合的 实时人体背部穴位识别

黄凌风 杨世龙 谢耀钦\*

(中国科学院深圳先进技术研究院 深圳 518055)

**摘要** 针灸是中医学的重要组成部分, 在世界各地均具有广泛的应用。然而, 传统针灸疗法的穴位定位依赖医生经验, 缺乏标准化, 导致疗效再现性较差, 阻碍了其推广。针灸机器人是一种智能医疗设备, 为针灸技术的标准化和推广提供了新契机。该文提出一种改进的 YOLOv8-Pose 模型——YOLO-PointMap, 旨在解决穴位密集分布和特征不明显等问题。通过引入动态卷积优化 C2f 模块和基于通道注意力的特征融合模块, 该文提出的模型在多尺度特征提取和融合方面的性能显著提升。实验结果表明, YOLO-PointMap 在测试集上的端点误差、正确关键点百分比和基于 COCO 标准的  $mAP_{50-95(Pose)}$  等指标优于现有模型, 其值分别达到了 3.27、1.00 和 84.90%, 尤其是在密集关键点检测和弱特征区域定位方面。这不仅为针灸机器人技术的发展提供了有力支持, 而且展现了 YOLO-PointMap 在虚拟现实和智能交互等领域的潜在应用价值。

**关键词** 穴位定位; 关键点检测; YOLO-PointMap; 动态卷积; 特征融合

中图分类号 TP399; R319 文献标志码 A doi: 10.12146/j.issn.2095-3135.20241129001

CSTR: 32239.14.j.issn.2095-3135.20241129001

### YOLO-PointMap: Real-time Human Back Acupoint Recognition Based on Lightweight Dynamic Feature Fusion

HUANG Lingfeng YANG Shilong XIE Yaoqin\*

(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

\*Corresponding Author: yq.xie@siat.ac.cn

**Abstract** As a vital part of traditional Chinese medicine, acupuncture has extensive application value all over the world. However, the reliance on practitioners' experience for acupoint localization in traditional acupuncture methods leads to a lack of standardization, restricting its reproducibility and broader adoption.

收稿日期: 2024-11-29 修回日期: 2024-12-17

基金项目: 国家自然科学基金项目 (U20A20373); 国家科技攻关计划项目 (2022YFC2409000); 深圳市科技计划项目 (KQTD20180411185028798, ZDSY20130401165820357); 介入手术机器人诊疗关键技术深圳市工程实验室项目 (XMHT20220104009)

作者简介: 黄凌风, 硕士, 研究方向为深度学习与人工智能、图像处理; 杨世龙, 硕士, 研究方向为深度学习与人工智能、图像处理; 谢耀钦 (通讯作者), 研究员, 研究方向为图像引导放射治疗、医学影像处理和分析、医学物理, E-mail: yq.xie@siat.ac.cn.

Acupuncture robots, as a kind of intelligent medical devices, offer new opportunities for standardizing and promoting acupuncture techniques. This paper introduces an improved YOLOv8-Pose model, YOLO-PointMap, designed to address challenges in dense acupoint distribution and weak feature recognition. By incorporating dynamic convolution to optimize the C2f module and introducing a channel-attention-based feature fusion module, the model achieves significant advancements in multi-scale feature extraction and integration. Experimental results show that the end point error (EPE), percentage of correct keypoints (PCK) and  $mAP_{50-95(Pose)}$  indexes of YOLO-PointMap on the test set are superior to the existing models, with the values reaching 3.27, 1.00 and 84.90% respectively, especially in dense key point identification and weak feature region localization. It provides strong support for the development of acupuncture robot technology, and shows the potential application value in the fields of virtual reality and intelligent interaction.

**Keywords** acupoint localization; keypoint detection; YOLO-PointMap; dynamic convolution; feature fusion

**Funding** This work is supported by National Natural Science Foundation of China (U20A20373); National Key Research and Development Program of China (2022YFC2409000); Shenzhen Science and Technology Program (KQTD20180411185028798, ZDSY20130401165820357); Shenzhen Engineering Laboratory for Diagnosis & Treatment Key Technologies of Interventional Surgical Robots (XMHT20220104009)

**Ethical Statement** The study was approved by the Institutional Review Board of the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (ID number: SIAT-IRB-230715-H0666), and written informed consent was obtained from each participant. It was also approved by the Ethics Committee of Xin-Huangpu Joint Innovation Institute of Chinese Medicine (ID number: 2023EC001), for studies involving humans

## 1 引言

随着人工智能和机器人技术的飞速发展, 智能化医疗设备在医疗领域的应用潜力越发凸显。在众多传统医学领域, 针灸历史悠久、疗效显著, 是中医的重要组成部分, 不仅在中国得到广泛应用, 在西方国家也被用于治疗各种慢性疾病, 如呕吐、热潮红、偏头痛和膝痛<sup>[1]</sup>。在常规治疗无效的情况下, 许多患者选择针灸缓解症状<sup>[2]</sup>。现代学者利用神经科学解释针灸的临床适用性, 认为它是炎症和疼痛的一种替代疗法<sup>[3-5]</sup>。一些研究表明, 针灸可以独立激活人体的自愈机制, 恢复体内平衡, 有效治疗免疫性疾病, 并被用于治疗抑郁症和帕金森病等精神障碍<sup>[6-9]</sup>。然

而, 传统针灸疗法的推广面临诸多挑战, 例如, 针灸点的定位高度依赖医生的经验, 缺乏统一标准。这种主观性不仅限制了疗效的再现性, 还阻碍了针灸在全球范围内的推广和普及<sup>[10]</sup>。

针灸机器人作为智能化医疗设备的重要方向, 为针灸技术的标准化和推广带来了新契机。针灸机器人能利用计算机视觉和人工智能技术实现穴位的自动化定位, 提高针灸治疗的精确性和效率。穴位定位任务本质上是一种特殊的关键点检测任务<sup>[11]</sup>, 与人体姿态估计中的关键点检测相比, 针灸点数量更多、分布更密集, 且往往缺乏明显的纹理特征。这些特点显著增加了特征提取的难度, 使穴位定位成为计算机视觉领域的一项极具挑战性的课题。值得注意的是, 密集关键点

检测不仅对针灸机器人至关重要，在虚拟现实和数字人等领域也具有广泛的应用潜力。例如，通过提升虚拟环境的沉浸感和数字人动作捕捉的精度推动智能交互技术的发展。

当前主流的目标检测方法(如YOLO系列模型)的检测速度和效率较优<sup>[12-14]</sup>。以YOLOv8-Pose为代表的模型结合了目标检测和关键点检测的能力，在人体姿态估计等任务中表现出色<sup>[15]</sup>。然而，由于下采样过程导致特征丢失，以及模型对全局特征的建模能力不足，因此这类模型在复杂任务(如针灸点定位)中仍存在局限性。例如，其颈部网络(Neck)部分对不同尺度特征的融合不够充分，难以有效捕捉针灸点定位所需的局部和全局信息。

针对上述问题，本研究创新性改进了YOLOv8-Pose模型，主要内容如下：

(1) 借鉴动态卷积<sup>[16]</sup>的思想，结合轻量化卷积的方法优化了C2f模块，在降低模型复杂度的同时提升特征提取的精准度。

(2) 设计、引入了基于通道注意力的特征融合模块，并替换了原模型的Neck部分。该模块能更有效地整合不同尺度的特征信息，增强对目标特征的捕获能力，从而提升针灸点的检测精度和鲁棒性。

(3) 在一个标定人体背部82个穴位的私有数据集上，就准确率而言，该模型的准确率、精度和效率都达到了先进(state of the arts, SOAT)水平。

## 2 相关工作

### 2.1 目标检测

目标检测技术经历了从传统方法到深度学习的快速演进。早期的传统方法(如Viola-Jones算法<sup>[17]</sup>和方向梯度直方图<sup>[18]</sup>)为目标检测奠定了基础。Viola-Jones算法利用Haar特征快速描述和提取人脸特征，并借助AdaBoost分类器进行特

征选择和构建强分类器，最终形成级联结构以高效准确地实现人脸检测；方向梯度直方图特征可捕捉图像的梯度方向信息，常用于行人检测。然而，这些方法在复杂场景下的检测性能仍存在局限性。

随着深度学习的兴起，卷积神经网络(convolutional neural networks, CNN)在目标检测中展现出强大的特征提取能力。双阶段检测方法，如R-CNN<sup>[19]</sup>系列，通过先生成候选区域，然后对这些区域进行精确的分类和边界框的回归优化，从而显著提高检测精度。其中，Faster R-CNN<sup>[20]</sup>引入了区域建议网络，实现了端到端的训练。然而，双阶段检测方法的计算开销较大，难以满足实时应用的需求。

单阶段检测方法，如YOLO系列，直接在整幅图像上同时完成目标检测和分类任务，无须生成候选区域。YOLO的端到端架构较简单，检测效率高，因此适用于实时性要求较高的任务。早期版本虽然在小目标检测上存在不足，但随着模型的不断改进，YOLO的检测精度显著提升。

### 2.2 关键点检测

关键点检测旨在定位物体上的重要特征点，如人体的关节点和面部特征点等。通常，关键点检测依赖目标检测提供的初始定位框架，即先通过目标检测确定物体的大致区域，再在该区域进行关键点检测。例如，在人体姿态估计中，关键点检测先通过目标检测确定人体的边界框，再在该区域定位关节点。

根据处理流程，关键点检测方法主要分为自顶向下和自底向上两类。

自顶向下方法先通过目标检测确定物体的边界框，然后在该区域定位关键点。Mask R-CNN<sup>[21]</sup>是这一方法的代表，它在Faster R-CNN的基础上增加了实例分割和关键点检测，实现了多任务的联合训练。然而，随着阶段目标检测技术的不断发展，YOLO不仅在速度上显著优于

Mask R-CNN, 而且在改进后, YOLO 的精度也逐渐追赶, 甚至超越了 R-CNN 家族<sup>[22]</sup>。

自底向上方法直接在图像中检测所有关键点, 通过后处理算法将关键点组合成具体的物体。OpenPose<sup>[23]</sup>是该方法的代表性算法, 它通过部分亲和场连接检测到的关键点, 实现多人的姿态估计。自底向上方法在复杂场景下能检测多个物体的关键点, 但对个体关键点的精度略低, 且需要更多的后处理步骤, 计算复杂度较高, 实时性不如自顶向下方法。

### 2.3 穴位定位

穴位定位是针灸疗法的关键环节, 传统上依赖医生的经验和解剖知识。早期研究尝试利用光学仪器、皮肤阻抗等方法辅助穴位定位。例如: 李延泽等<sup>[24]</sup>发明了简易的人体穴位识别仪; 何莉<sup>[25]</sup>设计了针对头颈部的穴位测量与定位系统; 朱明星<sup>[26]</sup>基于皮肤阻抗差异开发了穴位检测仪。这些方法在一定程度上提高了穴位定位的客观性, 但在精度和实用性方面仍存在限制。

随着数字化技术的发展, 研究者提出了“数字化经络”的概念, 并通过分形几何理论实现经络的数字化表示<sup>[27]</sup>。此外, 基于人体模板的按摩机器人、近红外光谱技术、计算机断层扫描 (computed tomography, CT) 和磁共振成像 (magnetic resonance imaging, MRI) 等手段也被应用于穴位识别和校准。例如: Litscher<sup>[28]</sup>利用近红外光谱技术识别头部穴位, 并结合多普勒超声和 MRI 验证其效果; 董诗绘等<sup>[29]</sup>基于工业机器人, 结合光学传感器、CT 和 MRI 对穴位进行识别和校准; 王聪<sup>[30]</sup>使用多个深度相机, 采用先整体后局部的方法, 即先使用全局相机识别人体姿态, 获取肢体位置, 再使用局部相机研究一条经络上的穴位定位, 最后解决肢体运动后穴位的重定位问题, 使激光针灸机器人得到三维坐标, 以完成后续治疗任务。

近年来, 虚拟现实 (virtual reality, VR) 和增

强现实 (augmented reality, AR) 技术被引入穴位定位领域。于俊伟等<sup>[31]</sup>在 iOS 系统上开发了耳针 AR 系统, 可辅助针灸教学和研究。杨星月等<sup>[32]</sup>建立了高精度的三维针灸数字人体模型, 结合 VR 设备, 能精确定位人体四肢上的穴位。这些方法虽然在虚拟环境中实现了高精度的穴位定位, 但在真实人体上的应用仍面临挑战。

## 3 方法

### 3.1 YOLO-PointMap 模型

针对针灸点密集分布和纹理特征弱的特点, 本研究基于 YOLOv8 -Pose 提出了 YOLO-PointMap 模型 (简称“YOLO-PM”)。YOLO-PM 由输入层 (Input)、主干网络 (Backbone)、颈部网络 (Neck) 和头部网络 (Head) 组成, 其结构如图 1 所示。YOLO-PM 模型的设计重点优化了特征提取和融合过程。

#### 3.1.1 YOLOv8-Pose 模型概述

YOLOv8-Pose 模型专注于关键点检测任务, 具有轻量化和高效性特征, 是本研究改进模型的基础。其结构包括输入层、主干网络、颈部网络和头部网络 4 部分。

(1) 输入层: 负责接收原始图像  $I_{\text{raw}}$ , 通过归一化、尺寸调整和马赛克数据增强<sup>[22]</sup>等预处理手段 (Preprocess) 生成输入给主干网络的图像  $I_{\text{in}}$ , 可提高模型对不同场景的鲁棒性。

$$I_{\text{in}} = \text{Preprocess}(I_{\text{raw}}) \quad (1)$$

(2) 主干网络: 采用 CSPNet<sup>[33]</sup>结构, 结合 C2f 模块提取多尺度特征  $F_{\text{backbone}}$ , 包括低层细节信息和高层语义信息。

$$F_{\text{backbone}} = \text{Backbone}(I_{\text{in}}) \quad (2)$$

(3) 颈部网络: 通过 PAN<sup>[34]</sup>-FPN<sup>[35]</sup>结构融合多尺度特征, 增强模型对大小目标的捕捉能力, 生成融合特征  $F_{\text{fused}}$ 。

$$F_{\text{fused}} = \text{PAN-FPN}(F_{\text{backbone}}) \quad (3)$$

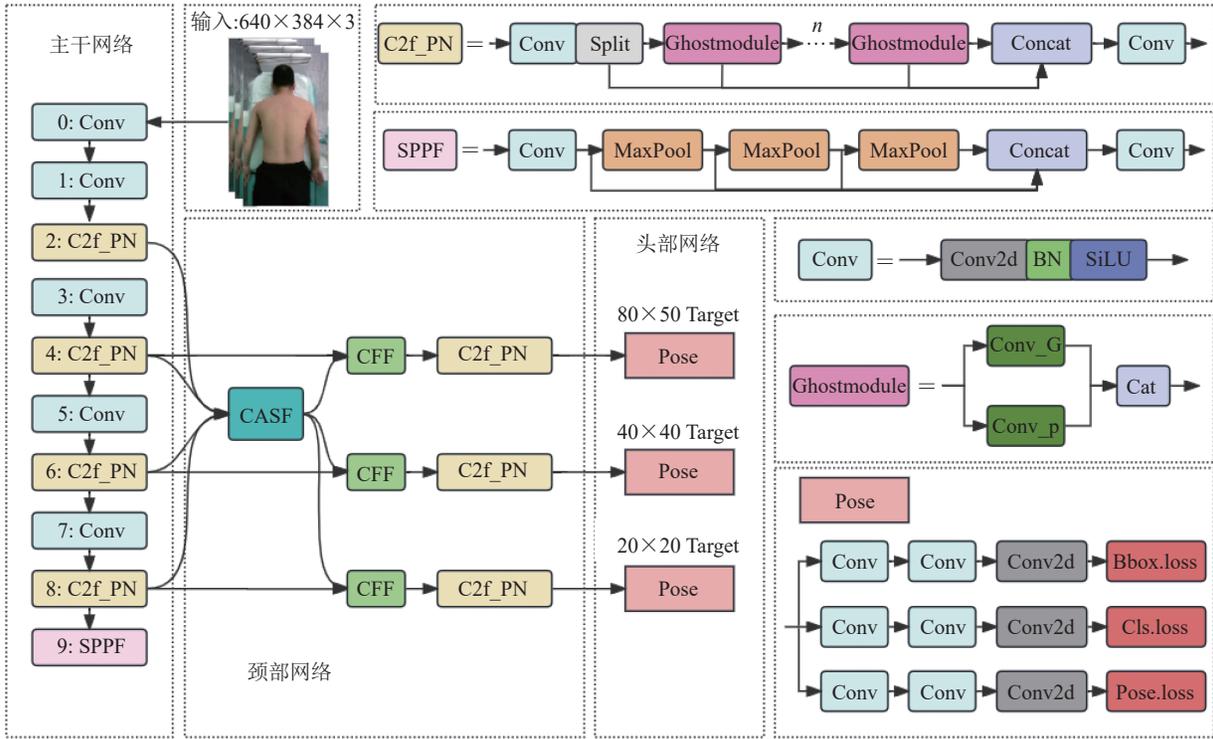


图1 YOLO-PointMap 网络结构

Fig. 1 YOLO-PointMap network framework

(4) 头部网络：预测目标的关键点位置，并得到每个关键点对应的置信度，然后通过计算分类损失、边界框回归损失和关键点损失，综合优化模型性能。这一系列的损失计算过程用 Head 表示，最终得到关键点的预测值  $\hat{P}$ 。

$$\hat{P} = \text{Head}(\mathbf{F}_{\text{fused}}) \quad (4)$$

### 3.1.2 轻量化的特征提取模块 C2f\_PN

本研究改进了原始的 C2f 模块，使用动态卷积和轻量化设计的 Ghost 模块，能使模型高效提取多尺度特征，显著降低了计算复杂度。

(1) 动态卷积：根据输入特征动态生成卷积核权重，能让模型针对不同输入特征适配最佳的卷积核。其核心思想是在一个卷积层中使用多个并行卷积核，通过一个路由网络生成权重，动态控制每个卷积核的参与程度。动态卷积的输出表示如下：

$$y = \sum_{i=1}^n \alpha_i \cdot f_i(x) \quad (5)$$

其中， $f_i(x)$  为第  $i$  个卷积核的输出； $\alpha_i$  为由输入特征通过路由网络生成的权重； $y$  为当前输入特征经过动态卷积层后的最终输出结果。路由网络对输入特征进行全局平均池化后，通过全连接层计算得到卷积核的权重。动态卷积可灵活调整卷积核的作用范围，增强模型对不同场景的适应能力，在处理细粒度特征分布（如针灸点）时表现优异。

(2) Ghost 模块：一种轻量化的卷积模块，旨在通过高效操作生成完整的特征图。其由两部分组成：第一部分通过标准卷积生成主特征 (Conv\_P)，这些特征是基础且高质量的；第二部分通过轻量化操作从主特征生成冗余特征 (Conv\_G)，用于补充主特征的多样性。主特征和冗余特征沿通道维度拼接，形成完整的输出特征图，并根据目标通道数进行裁剪。通过这种设计，Ghost 模块能以更低的计算和参数成本高效生成特征。

### 3.1.3 改进的特征融合颈部网络

为解决针灸点定位任务中多尺度特征的融合难题, 同时增强模型对关键点区域的关注和上下文理解能力, 本研究设计了基于通道注意力的选择性特征融合 (channel attention select fusion, CASF) 和上下文特征融合 (contextual feature fusion, CFF)。这两种策略在不同层面优化了多尺度特征融合的效率 and 表现。

(1) CASF 利用通道注意力机制动态选择最相关的特征通道, 以消除冗余信息, 并强化对关键点的关注。首先, 对输入特征图的每个通道进行全局平均池化, 生成全局描述向量; 然后, 通过全连接层生成每个通道的权重。权重反映了通道的重要性, 高权重的通道被保留, 并被用于后续的特征融合。通道选择的过程表示如下:

$$w_c = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot z)) \quad (6)$$

其中,  $z$  为通过全局平均池化生成的通道描述向量;  $W_1$  和  $W_2$  为全连接层的学习参数;  $\sigma$  为 Sigmoid 激活函数;  $w_c$  为被选择的特征通道组成的特征图。被选择的特征通道通过上下采样操作与其他尺度的特征进行尺寸对齐, 随后进行逐通道加权叠加, 实现多尺度特征的无缝融合。具体操作以 CASF 第一次处理为例: 将 Backbone 第 2、6、8 层的特征进行融合, 并与第 4 层对齐, 再传给 CFF 模块进行上下文尺度的融合。这种策略能显著增强模型对关键点区域的语义感知, 在密集关键点区域表现出色。

(2) CFF 通过结合局部特征和全局特征提升模型对关键点及其周围上下文信息的理解。其中, 局部特征源自较低层次的特征图, 蕴含丰富的细节信息, 为模型提供了关键点的精细化描述。而全局特征则通过对高层特征图进行全局平均池化生成, 其核心作用在于捕获关键点间的全局关联性, 从而使模型能洞察关键点间的宏观联系和整体协同性, 进而更全面、深入地把握关键点在全局范围内的相互作用和依存关系, 使模型

对关键点及其上下文信息的理解更精准透彻。在局部特征与全局特征的融合过程中, 全局特征被激活函数 (如 Sigmoid) 生成的权重加权, 以动态调整局部特征的贡献。特征交互过程可表示如下:

$$y_c = \text{Concat}(\mathbf{x}_{\text{local}}, \mathbf{x}_{\text{global}} \cdot \sigma(W \cdot \mathbf{x}_{\text{global}})) \quad (7)$$

其中,  $\mathbf{x}_{\text{local}}$  为局部特征;  $\mathbf{x}_{\text{global}}$  为全局特征;  $W$  为学习权重;  $y_c$  为经 CFF 模块处理后输出的特征图。CFF 通过将局部特征和加权后的全局特征进行拼接, 形成上下文增强的特征表示, 可有效建立关键点与周围区域间的语义关联, 在背景复杂或低对比度的场景下优势显著。

## 3.2 数据集与实验环境

本文数据集由本研究团队与深圳大学总医院的中医专家合作采集, 涵盖背部 84 个分布于 4 条经络的针灸穴位。数据采集涉及 210 名健康志愿者, 其中男性 155 名, 女性 55 名。入选标准如下: 年龄在 18 岁以上, 60 岁以下; 无背部皮肤病或伤口; 无背部畸形或手术痕迹; 无心脏病或其他严重疾病; 未怀孕且不在哺乳期。志愿者在采集前签署知情同意书, 并提供基本信息 (姓名、性别、年龄、身高和体重), 然后通过式 (8) 计算了身体质量指数 (body mass index, BMI), 结果如表 1 所示, 反映了数据集中 BMI 分布的全面性和多样性。

$$\text{BMI} = \frac{\text{weight}}{(\text{height})^2} \quad (8)$$

其中,  $\text{weight}$  为体重;  $\text{height}$  为身高。

表 1 数据集中各身体质量指数的比例

Table 1 The proportion of each body mass index in the dataset

| 体重 | 占比   |
|----|------|
| 过低 | 0.05 |
| 正常 | 0.77 |
| 超重 | 0.16 |
| 肥胖 | 0.02 |



如下:

$$mAP_{50-95(Pose)} = \frac{1}{10} \sum_{t=0.50}^{0.95} AP(t) \quad (11)$$

其中,  $AP(t)$  为阈值  $t$  下的平均精度。 $mAP_{50-95(Pose)}$  既能反映模型在低阈值下的整体定位能力, 又能评估模型在高阈值下的精细化检测表现。

本研究利用 EPE、PCK 和  $mAP_{50-95(Pose)}$  等 3 个指标对模型的关键点检测性能进行了定位偏差、准确性和综合精度等分析。

## 4 结 果

### 4.1 测试集的预测结果

为验证 YOLO-PM 在实际应用中的性能, 本

研究选取测试集上的部分样本进行预测结果展示。两例男性和两例女性的关键点预测如图 3 所示, 其中预测结果为红色点, 金标准为绿色点。

在穴位点密集分布的情况下, YOLO-PM 仍能精准捕获所有关键点的位置。测试结果表明, 模型在不同类型的样本中均能高精度定位穴位, 各关键点的置信度接近 1.0, 可靠性和稳定性较优。在穴位点的分布中, 部分关键点位于人体边缘或纹理特征较弱的区域, 这通常是传统方法容易失误的地方。YOLOv8 与 YOLO-PM 的预测结果如图 4 所示, 可以看到, YOLOv8 的肩部穴位预测结果与金标准偏差较大, 且右肩臑俞穴预测错误。YOLO-PM 在这些关键点的定位上表现出色, 预测结果与真实标注位置吻合。

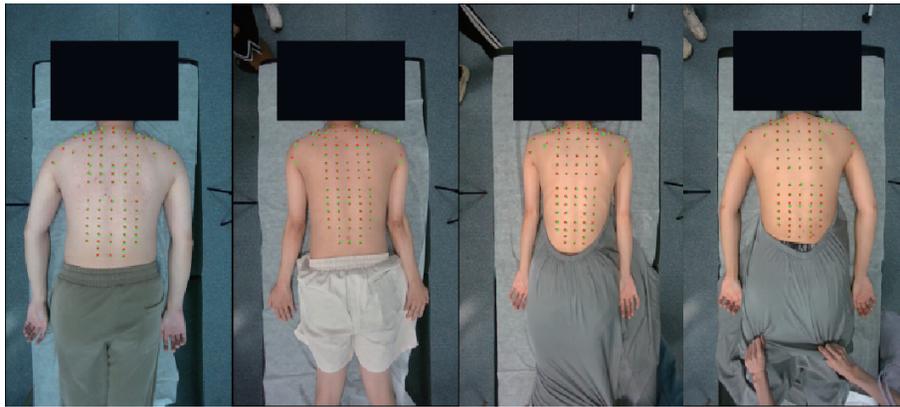


图 3 YOLO-PointMap 测试集预测结果

Fig. 3 Prediction results of YOLO-PointMap test set

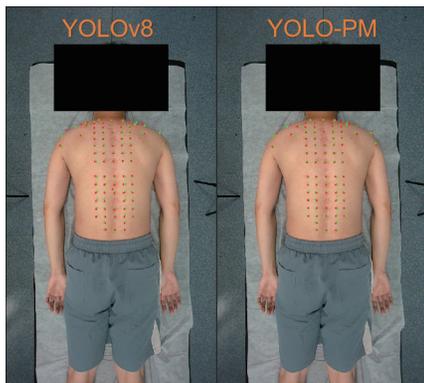


图 4 YOLOv8 与 YOLO-PM 的预测结果

Fig. 4 Prediction results of YOLOv8 and YOLO-PM

### 4.2 与现有方法的比较

本研究将 YOLO-PM 与 PVT<sup>[36]</sup>、Swin-pose<sup>[37]</sup>、RTMPose<sup>[38]</sup>和 ViTPose<sup>[39]</sup>进行了对比, 对比结果如表 2 所示。

(1) 精度分析: YOLO-PM 的 EPE 为 3.27, 远低于其他模型, 表明其在针灸点定位任务中误差最小, 定位精度最高。YOLO-PM 的 PCK 为 1.00, 大于其他模型, 表明其关键点检测的准确性较优。相比之下, PVT 和 Swin-pose 的 PCK 分别为 0.80 和 0.79, 表现较弱, 而 RTMPose 和

表 2 不同算法的对比

Table 2 Comparison of different algorithms

| 算法        | EPE   | PCK  | mAP <sub>50-95(Pose)</sub> /% | FLOPs/G | Params/M |
|-----------|-------|------|-------------------------------|---------|----------|
| YOLO-PM   | 3.27  | 1.00 | 84.90                         | 6.06    | 5.46     |
| PVT       | 10.75 | 0.80 | 78.42                         | 4.89    | 28.04    |
| Swin-pose | 10.21 | 0.79 | 69.62                         | 6.36    | 32.69    |
| RTMPose   | 11.23 | 0.89 | 82.12                         | 2.21    | 16.41    |
| ViTPose   | 10.03 | 0.91 | 82.06                         | 18.62   | 89.32    |

ViTPose 分别为 0.89 和 0.91，虽有较好表现，但仍不及 YOLO-PM。YOLO-PM 的 mAP<sub>50-95(Pose)</sub> 为 84.90%，优于其他模型，表明其定位精度最高。

(2) 复杂度分析：在深度学习模型的评估中，每秒浮点运算次数 (floating point operations per second, FLOPs) 和参数量 (Params) 是两个关键指标。FLOPs 衡量了模型在进行一次前向传播时所需的计算量，单位通常是 G (十亿次浮点运算)。参数量则表示模型中可训练参数的总数，单位通常是 M (百万个参数)。这两个指标分别反映了模型的计算复杂度和规模。较低的 FLOPs 表明模型在推理时更高效，而较少的参数量则通常表明模型更轻量化。YOLO-PM 的 FLOPs 为 6.06 G，与 Swin-pose (6.36 G) 接近，但显著低于 ViTPose 的 18.62 G。尽管比 RTMPose 的 2.21 G 稍高，但考虑到其显著的精度优势，这一复杂度是合理的。YOLO-PM 的参数量为 5.46 M，在所有对比模型中最小，仅为 ViTPose (89.32 M) 和 Swin-pose (32.69 M) 的极小部分，同时显著低于 PVT (28.04 M) 和 RTMPose (16.41 M)。这充分体现了 YOLO-PM 的轻量化设计优势。

### 4.3 消融实验

为评估 YOLO-PM 改进中各模块的贡献，本研究对比了 YOLOv8、改进的 YOLOv8 (更换 Neck 模块) 和最终的 YOLO-PM 模型的性能，通过逐步替换模块分析每次改进对性能的影响。验证集上的 mAP<sub>50-95(Pose)</sub> 收敛曲线如图 5 所示 (模型统一训练 2 000 epoch)，各模型的性能如表 3 所示。

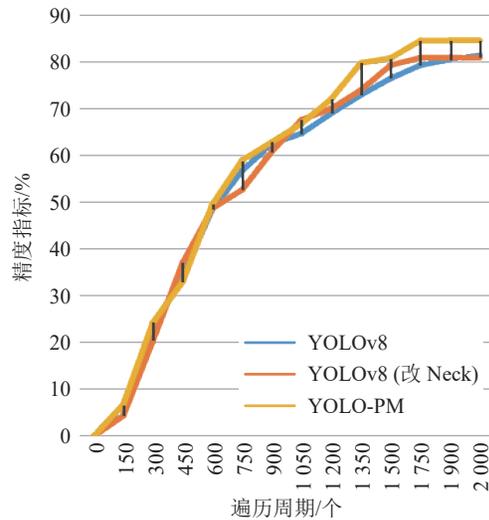


图 5 验证集平均精度曲线

Fig. 5 Average precision curve of the validation sets

表 3 消融实验结果

Table 3 Ablation study results

| 算法                  | EPE  | PCK  | mAP <sub>50-95(Pose)</sub> /% | FLOPs/G | Params/M |
|---------------------|------|------|-------------------------------|---------|----------|
| YOLOv8              | 8.34 | 0.98 | 81.54                         | 6.47    | 5.82     |
| YOLOv8 (更换 Neck 模块) | 5.46 | 0.99 | 81.03                         | 6.70    | 6.32     |
| YOLO-PM             | 3.27 | 1.00 | 84.90                         | 6.06    | 5.46     |

### (1) 验证集收敛曲线分析。

由图 5 可知, 3 种模型在验证集上的  $mAP_{50-95(Pose)}$  收敛曲线差异显著。

YOLOv8: 基础模型在训练初期的收敛速度较快, 但在后期趋于平稳, 最终达到 81.54% 的  $mAP_{50-95(Pose)}$ , 表明其对针灸点的检测具有一定能力, 但在密集关键点检测上仍有提升空间。

改进的 YOLOv8 (更换 Neck 模块): 替换 Neck 模块后, 模型的收敛速度略有提升, 最终在验证集上达到了 81.03% 的  $mAP_{50-95(Pose)}$ 。虽然整体精度变化较小, 但模型对细粒度特征的捕捉能力有所增强, 在早期训练阶段收敛更快。

YOLO-PM: 最终模型在验证集上的表现显著优于其他两种模型, 收敛速度最快, 并在后期达到了 84.90% 的  $mAP_{50-95(Pose)}$ , 表明动态卷积、通道注意力和上下文融合模块的引入有效提升了模型对多尺度特征的处理能力, 进一步优化了针灸点的精准定位。

### (2) 数值性能对比。

YOLOv8 的 EPE 为 8.34, PCK 为 0.98, 表明基础模型具有一定的定位误差; 改进的 YOLOv8 (更换 Neck 模块) 的 EPE 降至 5.46, PCK 提至 0.99, 但  $mAP_{50-95(Pose)}$  略有下降, 表明单一模块改进对整体性能的提升有限; YOLO-PM 在所有指标上均表现最优, EPE 降至 3.27, PCK 提升至 1.00,  $mAP_{50-95(Pose)}$  提升至 84.90%, 表明模型的综合性能得到提升。

## 5 总 结

本研究针对穴位定位任务提出了 YOLO-PM 模型, 通过改进 YOLOv8-Pose 的关键模块, 显著提升了模型的检测精度和鲁棒性。实验结果表明, 与现有主流方法相比, YOLO-PM 在测试集上的 EPE、PCK 和  $mAP_{50-95(Pose)}$  等均最优。通过

引入动态卷积和上下文特征融合模块, 本文模型克服了传统方法在特征提取和多尺度融合上的不足, 为中医穴位定位的智能化和标准化提供了技术支持。

YOLO-PM 的实验效果虽然较好, 但在实际应用中, 仍有一些方向值得进一步探索。

(1) 模型的通用性与扩展性: 针灸点的定位精度在其他部位 (如四肢或头部) 上的表现尚未验证, 未来可以扩展到全身的针灸点定位任务, 并适配更多人体形态和姿势。

(2) 实际临床验证: 将模型应用于真实针灸机器人, 结合不同体型、皮肤条件和光照环境进行更大规模的临床实验, 验证其在实际医疗场景中的适应性和可靠性。

未来, 随着 YOLO-PM 模型在轻量化、鲁棒性和多样化场景中的持续优化, 针灸机器人技术有望进一步发展, 从而推动传统针灸技术的全球化推广和智能化应用。传统针灸面临技能差异和个体化需求的挑战, 而 YOLO-PM 能精准定位针灸点, 减少误差, 确保治疗的一致性和安全性, 为全球推广提供可靠支持。特别地, 在不同文化和人体结构下, YOLO-PM 可以根据需求调整治疗方案, 提升针灸技术的普适性。

此外, YOLO-PM 的关键点检测能力在 VR 和动作捕捉领域的应用潜力较大。在 VR 培训中, 结合 YOLO-PM, 针灸机器人能提供更精确的模拟训练, 帮助医师提升技能。YOLO-PM 还可以在运动医学和康复领域中精准捕捉动作, 为智能交互和动作分析提供支持, 推动多个行业的发展。

为实现这些拓展应用, 须优化 YOLO-PM 的计算效率, 确保其在资源受限设备上的实时性和稳定性, 并根据不同应用场景进行定制训练。通过跨学科合作和技术迭代, YOLO-PM 能进一步推动智能医疗、虚拟现实等领域的创新, 拓展其在全球范围内的应用潜力。

## 参 考 文 献

- [1] Yang ES, Li PW, Nilius B, et al. Ancient Chinese medicine and mechanistic evidence of acupuncture physiology [J]. *Pflügers Archiv-European Journal of Physiology*, 2011, 462: 645-653.
- [2] Mao JJ, Zhu BY, Wang HS. Acupuncture-based treatments on cervical spondylosis of vertebral artery type [J]. *TMR Non-Drug Therapy*, 2021, 4(1): 6.
- [3] Epstein AS, Liou KT, Romero SAD, et al. Acupuncture vs massage for pain in patients living with advanced cancer: the IMPACT randomized clinical trial [J]. *JAMA Network Open*, 2023, 6(11): e2342482.
- [4] Shah S, Godhardt L, Spofford C. Acupuncture and postoperative pain reduction [J]. *Current Pain and Headache Reports*, 2022, 26(6): 453-458.
- [5] Oh JE, Kim SN. Anti-inflammatory effects of acupuncture at ST36 point: a literature review in animal studies [J]. *Frontiers in Immunology*, 2022, 12: 813748.
- [6] Wang M, Liu W, Ge J, et al. The immunomodulatory mechanisms for acupuncture practice [J]. *Frontiers in Immunology*, 2023, 14: 1147718.
- [7] Liu S, Wang Z, Su Y, et al. A neuroanatomical basis for electroacupuncture to drive the vagal-adrenal axis [J]. *Nature*, 2021, 598(7882): 641-645.
- [8] Lin JG, Kotha P, Chen YH. Understandings of acupuncture application and mechanisms [J]. *American Journal of Translational Research*, 2022, 14(3): 1469-1481.
- [9] Yang NN, Lin LL, Li YJ, et al. Potential mechanisms and clinical effectiveness of acupuncture in depression [J]. *Current Neuropharmacology*, 2022, 20(4): 738-750.
- [10] 武晓冬, 刘保延. 我国针灸标准化的现状及面临的挑战与对策 [J]. *中国针灸*, 2019, 39(4): 343-348.
- Wu XD, Liu BY. Present situation, challenge and strategy of acupuncture standardization in China [J]. *Chinese Acupuncture & Moxibustion*, 2019, 39(4): 343-348.
- [11] Li YL, Teng YS, Huang YQ, et al. AIR-Net: acupoint image registration network for automatic acupoint recognition and localization [J]. *Displays*, 2024, 83: 102743.
- [12] Han X, Chang J, Wang K. You only look once: unified, real-time object detection [J]. *Procedia Computer Science*, 2021, 183(1): 61-72.
- [13] Farhadi A, Redmon J. YOLOv3: an incremental improvement [Z/OL]. *arXiv Preprint, arXiv: 1804.02767*, 2018.
- [14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 7263-7271.
- [15] Maji D, Nagori S, Mathew M, et al. YOLO-pose: enhancing YOLO for multi person pose estimation using object keypoint similarity loss [C] // *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2022: 2637-2646.
- [16] Han K, Wang Y, Guo J, et al. ParameterNet: parameters are all you need for large-scale visual pretraining of mobile networks [C] // *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 15751-15761.
- [17] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [C] // *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, 1: I-511-I-518.
- [18] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // *Proceeding of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005: 886-893.
- [19] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and

- semantic segmentation [C] // Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [20] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [21] He K, Gkioxari G, Dollár P, et al. Mask R-CNN [C] // Proceedings of the 2017 IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [22] Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: optimal speed and accuracy of object detection [Z/OL]. arXiv Preprint, arXiv: 2004.10934, 2020.
- [23] Cao Z, Simon T, Wei SE, et al. Realtime multi-person 2D pose estimation using part affinity fields [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7291-7299.
- [24] 李延泽, 刘彦博, 朱凯晖, 等. 简易人体穴位识别治疗仪 [J]. 吉林大学学报 (信息科学版), 2010, 28(6): 609-615.  
Li YZ, Liu YB, Zhu KH, et al. Simple human body acupoint recognition therapy instrument [J]. Journal of Jilin University (Information Science Edition), 2010, 28(6): 609-615.
- [25] 何莉. 头颈部危险穴位测量与定位方法研究 [D]. 天津: 天津大学, 2011.  
He L. Study on measurement and location of dangerous acupoints in head and neck [D]. Tianjin: Tianjin University, 2011.
- [26] 朱明星. 中医腧穴智能定位系统设计 [D]. 长沙: 中南大学, 2012.  
Zhu MX. Design of intelligent positioning system for acupoints in traditional Chinese medicine [D]. Changsha: Central South University, 2012.
- [27] 徐天成, 曹佳楠, 王璇, 等. 试探使用分形理论实现经络数字化方法与意义 [J]. 辽宁中医药大学学报, 2016, 18(7): 136-141.  
Xu TC, Cao JN, Wang X, et al. Exploring the method and significance of using fractal theory to realize the digitization of meridians [J]. Journal of Liaoning University of Traditional Chinese Medicine, 2016, 18(7): 136-141.
- [28] Litscher G. 针灸现代化研究进展 [J]. 湖北中医药大学学报, 2015, 3: 1-6.  
Litscher G. Research progress on modernization of acupuncture and moxibustion [J]. Journal of Hubei University of Chinese Medicine, 2015, 3: 1-6.
- [29] 董诗绘, 王旭. 基于 ABB 机器人的人体经络腧穴识别 [J]. 现代制造技术与装备, 2018, (2): 187-188.  
Dong SH, Wang X. The human body meridians recognition based on ABB robot [J]. Modern Manufacturing Technology and Equipment, 2018, (2): 187-188.
- [30] 王聪. 激光针灸机器人视觉寻穴方法研究 [D]. 北京: 北京邮电大学, 2020.  
Wang C. Robot vision for laser acupuncture point method research [D]. Beijing: Beijing University of Posts and Telecommunications, 2020.
- [31] 于俊伟, 房露, 鲁传政. 耳部针灸模型的移动增强现实虚实融合 [J]. 计算机工程与设计, 2018, 39(5): 1497-1501.  
Yu JW, Fang L, Lu CZ. Virtual-real fusion of auricular acupuncture model in mobile augmented reality [J]. Computer Engineering and Design, 2018, 39(5): 1497-1501.
- [32] 杨星月, 许安萍, 张宇沁, 等. 虚拟仿真技术在经络腧穴学教学中的应用 [J]. 中国中医药现代远程教育, 2021, 19(12): 18-20.  
Yang XY, Xu AP, Zhang YQ, et al. Application of virtual simulation technology in the teaching of meridian acupoints [J]. Modern Distance Education of Chinese Medicine, 2021, 19(12): 18-20.
- [33] Wang CY, Liao HYM, Wu YH, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [34] Lin TY, Dollár P, Girshick R, et al. Feature pyra-

- mid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [35] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [36] Wang W, Xie E, Li X, et al. Pyramid vision Transformer: a versatile backbone for dense prediction without convolutions [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 568-578.
- [37] Xiong Z, Wang C, Li Y, et al. Swin-pose: Swin Transformer based human pose estimation [C] // Proceedings of the 2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval, 2022: 228-233.
- [38] Jiang T, Lu P, Zhang L, et al. RTMPose: real-time multi-person pose estimation based on MMPose [Z/OL]. arXiv Preprint, arXiv: 2303.07399, 2023.
- [39] Xu Y, Zhang J, Zhang Q, et al. ViTPose: simple vision ? Transformer baselines for human pose estimation [J]. Advances in Neural Information Processing Systems, 2022, 35: 38571-38584.