arXiv:2603.01147v1 [cs.CV] 1 Mar 2026

ConVibNet：基于频率启发特征的连续插入过程中的针检测

Jiamei Guo†, 1，Zhehao Duan†, 1，Maria Neiiendam[1,3]，
Dianye Huang*, 1, 2，Nassir Navab[1,2]，ZhongliangJiang[1,2,4]

[1]计算机辅助医疗操作与增强现实（CAMP），慕尼黑工业大学，德国慕尼黑。[2]慕尼黑机器学习中心（MCML），德国慕尼黑。[3]丹麦技术大学，丹麦孔恩斯灵比。[4]香港大学，中国香港。

†共同贡献。

*通讯作者。电子邮箱：dianye.huang@tum.de；贡献作者：jiamei.guo@tum.de；zhehao.duan@tum.de；maria.neiiendam@gmail.com；nassir.navab@tum.de；zljiang@hku.hk；

摘要

目的： 超声引导针介入广泛应用于临床实践，但其成功与否在很大程度上取决于针的精准放置，而超声图像中针的可见性较差且时有时无，常常对此造成阻碍。现有方法仍受伪影、遮挡和低对比度的限制，且往往无法支持实时连续插入。为克服这些挑战，本研究提出了一种用于连续针检测的稳健实时框架。

方法： 我们提出了 ConVibNet，这是 VibNet 的扩展，用于检测可见性显著降低的针，并解决插入过程中的实时连续针跟踪问题。ConVibNet 利用连续超声帧之间的时间依赖关系，在动态场景中连续估计针尖位置和针杆角度。为增强对针尖运动的时间感知能力，我们提出了一种新颖的交集与差分损失，显式利用连续帧之间的运动相关性。此外，我们还整理了一个专用数据集，用于模型开发与评估。

结果： 所提出的 ConVibNet 模型的性能在我们的数据集上进行了评估，相较于基线 VibNet 和 UNet-LSTM 模型展现出更高的精度。具体而言，ConVibNet 的尖端误差为 2.80±2.42 mm，角度误差为 1.69±2.00°。这些结果相当于 0.75

mm 的针尖定位精度提升（优于表现最佳的基线），同时保持了实时推理能力。

结论：ConVibNet 将时间相关性建模与新颖的交集与差分损失相结合，提升了超声引导介入中实时针检测的准确性和鲁棒性，并展现出集成到自主插入系统中的巨大潜力。

关键词：针跟踪、超声引导介入、针振动、超声图像分析

1 引言

超声引导针介入是活检、区域麻醉和消融等医疗操作中应用最广泛的技术之一，这得益于超声（US）成像的实时性和无辐射特性 [1]。然而，其疗效在很大程度上取决于对皮肤下插入针的位置进行精确定位。由于散斑噪声和低对比度的固有局限，超声图像中的自动针检测仍面临挑战，导致针可见性降低 [2]。超声图像狭窄的成像平面，以及针容易受到类针伪影遮挡的影响，使这一挑战进一步加剧，导致针间歇性可见，并给自动检测带来重大障碍 [3, 4]。

超声成像中的传统针检测方法主要依赖手工特征，例如统计滤波器 [5] 和Gabor滤波器 [6], ，并结合几何建模来利用针杆和针尖的独特外观。随着卷积神经网络（CNN）的最新进展，数据驱动方法展现出更强的鲁棒性和泛化能力 [7–9]。具有代表性的例子包括将全卷积网络与基于区域的卷积神经网络相结合以进行针轨迹估计 [10], ，以及在粗针活检过程中采用自适应压缩和基于 Radon 变换的分割来实现自动引导的方法 [11]。尽管这些研究取得了令人满意的结果，但仍面临挑战，尤其是依赖单帧信息时，容易受到遮挡、声学伪影和低对比度成像条件的影响。

为应对这些挑战，研究者将时间信息纳入检测框架，以利用运动线索并提升跟踪稳定性 [12]。代表性方法包括将 CNN 与长短期记忆（LSTM）模块相结合的架构 [13] ，以及加入受卡尔曼滤波启发的组件的编码器—解码器网络，以捕捉运动连续性并减少针尖定位误差 [4]。最近，张等 [3] 提出了 MambaXCTrack，该方法采用结构化状态空间建模，并结合互相关和隐式运动提示，实现超声引导下的针跟踪，有效减轻了图像噪声和针尖间歇性不可见问题。然而，如果推理阶段的背景纹理不同于训练期间的背景纹理，仅依赖针视觉外观随时间变化的方法，其泛化能力可能有限。

硬件增强方法包括在针表面添加凹槽、凹坑或专用涂层等特征，以提高其在超声成像中的可见性，例如使用回声增强针 [14], ，或使用具有多普勒闪烁特征的针 [15]。然而，这些方法会增加针具成本，还可能引入额外的成像伪影。为降低对输入空间纹理的依赖，一种可行策略是考察连续超声图像中像素强度的变化。Beigi 等人 [16] 探索了利用手动震颤运动辅助B型超声中的针检测；当针杆最初在穿刺部位附近可见时，该方法使用凸阵探头取得了更好的结果。然而，在体内试验中，由于周围组织产生的运动噪声，检测性能有所下降。为解决手持探头中的震颤诱发伪影，他们进一步提出了一种时空跟踪框架 [17], ，但该方法依赖手工特征和多尺度空间分解，因此处理速度仅为每秒一帧。最近，Huang等人 [2] 开发了 VibNet，通过使针产生机械振动来生成周期性运动特征，并通过频域特征提取捕捉这些特征。尽管即使针在B型超声成像中几乎不可见，VibNet 仍能提高检测鲁棒性，但它仅适用于静态针检测，无法处理连续插入。

为进一步应对这一挑战，本研究探索在实时插入过程中利用频率特征检测超声图像中的移动针。为更好地捕捉时间动态，我们提出了一种新颖的交集与差异损失，明确利用连续帧之间的运动相关性，从而增强时间感知能力和一致性。为缓解超声图像中近乎不可见针的标注难题，我们还开发了一套数据采集平台，配备 NDI跟踪系统和针导向器，用于自动标注，并结合数据质量检查，确保真值标签可靠。与基线方法的比较和消融实验结果表明，ConVibNet实现了卓越的实时性能，验证了其设计的有效性以及频域时间运动建模的优势。总而言之，本研究将 ConVibNet 定位为一种前景广阔的超声引导针检测框架，有望集成到电动插入系统和未来的自主平台中。

2 方法

本节介绍 ConVibNet，这是一种旨在捕捉连续超声帧间时间依赖关系、实现实时针检测的框架。具体而言，ConVibNet 将 VibNet 扩展至动态插入场景，可连续估计针尖位置和针杆角度。此任务不仅需要处理有意施加的小幅度振动，还要应对额外的插入诱发的针运动，而这些运动可能带来混杂效应。因此，一个关键问题随之而来：在插入条件下，频域信息是否仍足够显著？为探究这一问题，我们分析了以 30帧/秒 录制的超声序列，其中针按照图3（a）所示的机构振动。接下来

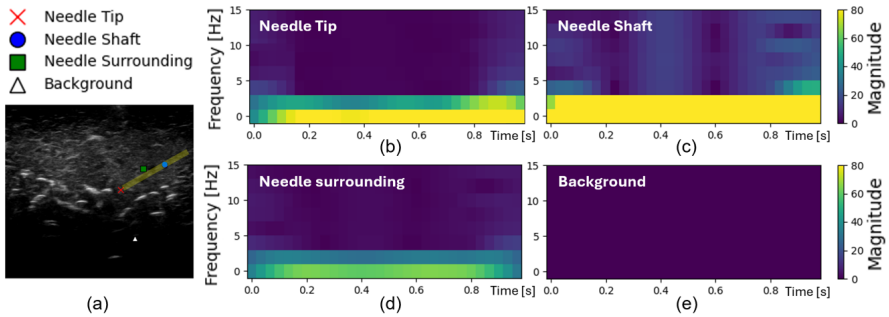


图1：美国图像序列的频率分析。(a) 第 t帧中选取用于分析的像素位置，黄色透明线表示针的插入角度为30°。(b)~(e) 对应于针尖、针杆、周围组织和背景的频谱图。

如图1所示，我们在一个包含 30 帧的序列中手动选取了四个代表性点：针尖、针杆、周围组织和背景。我们使用短时傅里叶变换（STFT）将这些点的时间强度变化转换为时频谱图。结果表明，与组织 [图 1（d）] 和背景 [图 1（e）], 相比，针尖 [图 1（b）]和针杆 [图 1（c）] 呈现出更强且更清晰的频率分量。频谱图的幅值随着与针杆距离的增加而逐渐减小，在周围组织中最弱。这些观察结果证实，频域特征沿针轨迹依然显著，直观地为连续针检测奠定了基础。

2.1 运动相关性的交集与差异损失

图 2 概述了所提出的方法。将VibNet扩展到连续场景会带来两大挑战。首先，实时针检测受到霍夫变换模块的阻碍，正如 [2], 中所报告的，该模块计算成本很高，因此不适合实时针尖跟踪。为解决这一问题，我们仅保留来自VibNet的运动编码器以及频率特征提取和聚合模块 [2], ，同时用分割头替换深度霍夫变换（DHT） [，见图2（a）]。其次，为缓解类别不平衡，我们采用焦点损失 [18], ，其定义见式（1）。该损失非常适合细针检测 [19]。它自适应地降低已正确分类样本的权重，减少大量负样本的影响并强调正样本，从而引导模型更有效地关注针掩膜。

$$\mathcal{L}_f = -\frac{1}{N}\sum_i \left[y_i \left(1-\hat{y}_i\right)^{\eta} \log\left(\hat{y}_i\right) + \left(1-y_i\right)\left(1-\hat{y}_i\right)^{\gamma} \hat{y}_i^{\eta} \log\left(1-\hat{y}_i\right) \right] \tag{1}$$

其中 $\eta = 2$ 和 $\gamma = 4$； y_i 和 $\hat{}y_i$ 分别表示真实标注和预测输出。

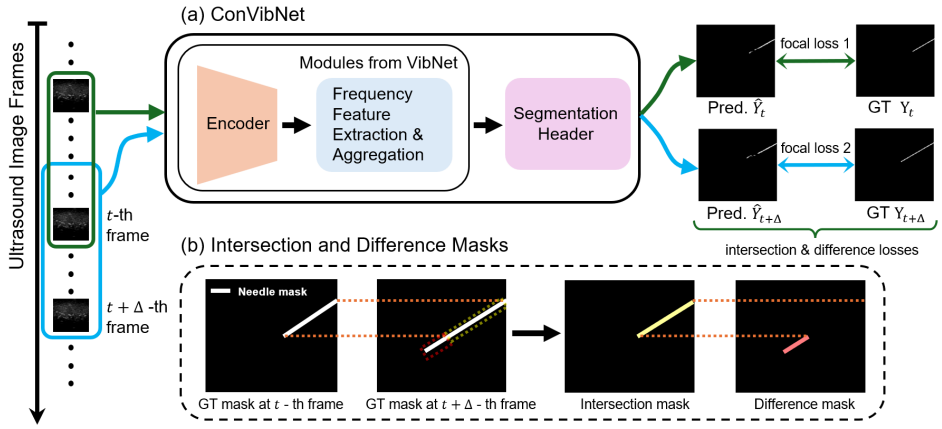


图 2：所提出方法的概览。(a) ConVibNet 的架构。(b) 交集掩码和差异掩码的计算示意图。

其次，VibNet 处理由 L 个连续帧组成的序列，但只生成一个分割输出 Y。相比之下，我们的任务涉及连续插入针具，需要针对每个序列的最后一帧进行分割。引入频域特征会带来时间聚合，使得难以提取仅与最后一帧相关的信息。为解决这一问题，我们扩展了损失函数，引导 ConVibNet 通过比较两个输入序列的预测结果来捕捉时间运动相关性，这两个序列的最后一帧相隔 Δ 个时间步。在我们的实现中，每个序列包含 $L = 30$ 帧，且 $\Delta = 5$。

如图2所示，两个序列由 ConVibNet 分别处理，并针对两个输出计算焦点损失 [见公式（1）]。随后，我们引入两个辅助损失项：交集损失（$\mathcal{L}_{\text{inter}}$）和差异损失（$\mathcal{L}_{\text{diff}}$）。交集损失通过强调一致的预测，旨在提高模型在细粒度区域的准确性；差异损失则促使模型捕捉序列之间的时间动态，从而增强其学习时间依赖特征的能力。图 2 (b) 展示了两个真实标注或预测分割输出之间交集掩码和差异掩码的计算方式。

ConVibNet 的总损失函数表示为四个组成部分的加权和：每个输入序列的焦点损失、交集损失和差异损失。

$$\mathcal{L} = \mathcal{L}_f^{(t)} + \mathcal{L}_f^{(t+\Delta)} + \alpha \mathcal{L}_{\text{inter}} + \beta \mathcal{L}_{\text{diff}} \tag{2}$$

其中

$$\mathcal{L}_{\text{inter}} = \text{BCE}\left(\hat{Y}_t \cdot \hat{Y}_{t+\Delta},\ Y_t \cdot Y_{t+\Delta}\right), \quad \mathcal{L}_{\text{diff}} = \text{BCE}\left(|\hat{Y}_t - \hat{Y}_{t+\Delta}|,\ |Y_t - Y_{t+\Delta}|\right),$$

α 和 β 分别是控制交集损失与差异损失贡献的超参数；“ $\cdot$ ”表示逐元素乘法。

值得注意的是，只有当模型达到相对稳定的状态后，才会在训练过程中引入差异损失。根据经验，过早应用该损失可能会妨碍收敛，这很可能是因为在模型尚未学会基本表征之前就强制要求动态一致性，会增加过多复杂性。在本研究中，从第二个训练轮次开始启用差异损失。模型采用Adam算法进行优化，学习率为 1×10^{-4}，批量大小为4。为防止过拟合，我们采用了早停策略；鉴于下文所述的数据集规模，训练通常在第 3^{rd} 个训练轮次前收敛，此时仅需对训练集进行少量完整遍历。由于预测得到的针分割掩膜可能不连续，我们在后处理中应用RANSAC算法拟合直线，以提高稳健性。最后，将距离图像底部最近的正像素投影到拟合直线上，以确定针尖位置。

2.2 数据准备

2.2.1 实验设置

超声成像的一大挑战是针的可见性较差，尤其是在组织穿刺过程中，因此难以直接将针尖标注为真实标注。为解决这一问题，我们搭建了一个数据采集平台 [图3(a)] ，使用离体猪组织模拟真实的软组织条件。为增加数据集的多样性，我们随机改变插入位置。将一根18号 $90\ mm$ 针沿导向器插入，插入角度为预设的两种角度之一（$\theta = 15^\circ$ 或 30°），以确保轨迹一致，涵盖浅角度插入和相对较陡的插入情形，同时满足实验设置的限制。插入过程中，步进电机（28BYJ-48-5V，PollinElectronicGmbH，德国）通过电机控制器驱动，使针杆以约 $2.5\ Hz$ 的频率振动，为后续分析引入时间频率分量。一个3D打印连接件将振动从电机轴传递至针座；安装在连接件上的两个被动标记则可借助NDI系统（Polaris Vicra，Northern Digital Inc.，加拿大）实时光学跟踪针的位移，采样频率为 $100\ Hz$。使用Siemens ACUSON Juniper系统（SIEMENS AG，德国）和线阵探头（12L3，探头接触面 $51.3\ mm$）采集超声图像，成像深度为4.5 cm。图像流通过Magewell视频采集卡（Magewell，中国）实时捕获并传输至工作站，以便进一步处理。

2.2.2 数据采集与标注流程

如图3(b)所示，每段插入视频的针掩膜标注分为**第 1 步**：将超声探头固定在机械臂上，沿预定义的针导向器插入针，并调整超声波束角度以获得最佳可视性。**第 2 步**：针尖可见后，在当前帧中手动标注针尖，随后重置波束角度。**第 3 步**：插入过程中，NDI系统跟踪针上的被动标记，并通过空间转换将其三维轨迹与每个超声帧同步 [20]。根据标记位置计算针尖在帧间的相对位移，从而推导出

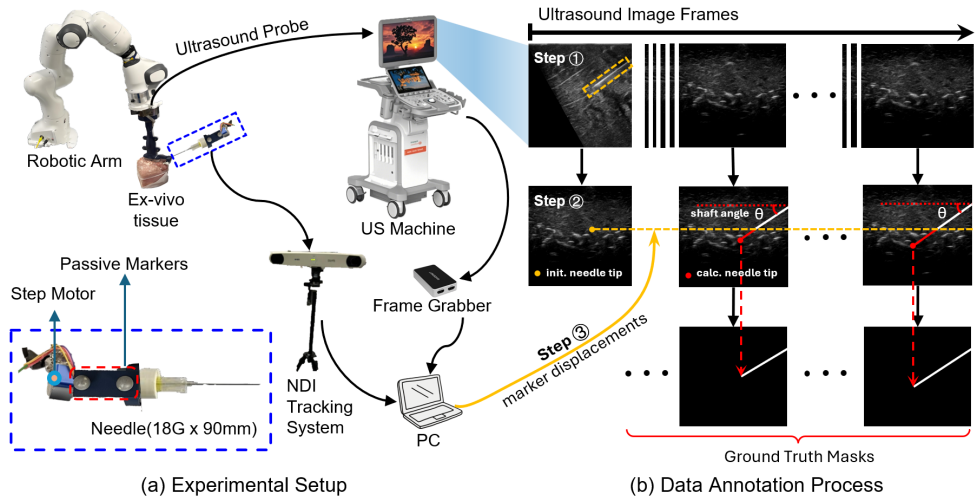


图 3：数据采集流程概览。(a) 硬件设置及连接（见 第 *2.2.1*节）。*(b)* 超声图像序列标注流程（见 第 *2.2.2*节）。

后续帧中的真实针尖位置，该位置以初始手动标注为基础得出。即使针在超声图像中不清晰，此过程仍能生成准确的针掩膜。

在每次插入试验中，针尖都会反复向前推进并部分回撤数次，以模拟典型的临床探查和调整操作。该过程持续进行，直至针完全拔出。对于每个预设插入角度，均重复该操作 60 次；每次试验的组织表面位置各不相同，共获得 120 个超声视频。在数据集整理过程中，**步骤 1**和**步骤 2**中超声波束尚未恢复至默认方向的帧及其对应掩膜均被剔除。尽管使用了针导向器，实际针路径仍可能因与软组织的机械相互作用而偏离预设轨迹；软组织的异质性和黏弹性会使针杆受到不均匀的作用力，从而导致弯曲或偏转 [21]。为考虑这一情况，我们进行了人工质量检查：对每个视频，都目视检查生成的掩膜，并将其与针的大致位置进行比较。尽管针在超声图像中几乎不可见，但可以根据周围组织的运动推断其大致位置。标注掩膜与这一粗略估计之间存在明显角度偏差的序列均被剔除。如表1所示，所收集的数据集包含 106 个视频，并按 80%、10% 和 10% 的比例划分为训练集、验证集和测试集。训练期间，我们对序列进行在线增强，包括水平翻转以及调整图像对比度和亮度，以提高鲁棒性，并缓解数据集规模有限所导致的过拟合。每个序列包含连续的 30 帧超声帧。

3 评估与结果

我们系统地将所提出的模型与动态针检测的相关基线方法进行评估。实验比较了传统的基于强度的方法与我们的频域变体在针尖定位和插入角度估计方面的表现。所有模型均在相同的数据集上训练，并使用以下三项指标评估性能：*(i)* 尖端误差（毫米）：预测针尖位置与真实针尖位置之间的欧几里得距离均值；*(ii)* 角度误差（°）：预测针杆角度与真实针杆角度之间的平均绝对偏差；*(iii)* 成功率（%）：尖端误差为 < 10 毫米且角度误差为 $< 15°$ 的测试样本所占比例，反映针检测的可靠性。

3.1 对比研究

为验证所提出的方法，我们将 ConVibNet 与两种基线方法进行比较：(i) VibNet w/o DHT，其中将原始 DHT模块替换为与 ConVibNet 相同的分割头，以确保公平比较，因为 DHT 的计算开销过大，不适用于实时应用；(ii) UNet-LSTM [13], ，它将 UNet 与 LSTM 相结合，以捕捉时间依赖关系。VibNet w/o DHT 作为频域基线，而 UNet-LSTM 则代表传统的基于强度的时空模型。尽管存在回声增强针 [14] 或基于多普勒的跟踪 [15] 等专用硬件解决方案，但它们往往成本较高，或需要特定的成像模式。如表 2所示，ConVibNet 的尖端误差最低，为 2.80±242 mm，比 VibNet w/oDHT 低 0.75 mm，比 UNet-LSTM 低 0.80 mm。虽然角度误差略微增至 1.69±2.00° ，高于两种基线的 1.59°，但 ConVibNet 仍保持了相当的角度精度。值得注意的是，其成功率达到 79.6%，而 VibNet w/o DHT 为 63.7%，表明其在动态针检测方面具有更高的整体可靠性。在推理方面，ConVibNet 在 RTX 1080Ti GPU 上的单样本推理时间为 33 ms，与临床超声成像中常用的 30 Hz 帧率相匹配。

图 4 中的定性结果展示了 ConVibNet 在不同插入角度下取得的改进。在 $30°$处，ConVibNet 生成了更连续的分割掩膜，能够准确捕捉针尖；而基线方法产生的针杆检测结果却支离破碎，导致尖端误差较大。在 $15°$处也观察到了类似的改进。该情况更具挑战性，因为针周围有外观类似针的细长组织结构。这些结构增加了运动模式歧义，进而降低了跟踪性能。在此情况下，VibNet 会产生过多且嘈杂的检测结果，经过后处理后对针长估计过长；而 UNet-LSTM 仅能检测到入针点附近的一小段，

表 1：精选数据集概览

插入角度	总计	训练集	验证集	测试集
15°	50/10380	40/8629	5/825	5/926
30°	56/22965	44/18157	6/2298	6/2510

数值以以下格式呈现： *num of videos/num of sequences*。

表 2：与基线模型相比的针检测性能

方法	尖端误差 (mm)	角度误差 (°)	成功率 (%)
VibNet [2] 不含 DHT	3.55 ± 2.65	1.59 ± 2.13	63.7
UNet-LSTM [13]	3.60 ± 2.75	$\mathbf{1.59 \pm 1.78}$	62.7
ConVibNet† （本文方法）	$\mathbf{2.80 \pm 2.42}$	1.69 ± 2.00	**79.6**

角度误差和尖端误差采用均值±标准差的格式表示。†所提出的 ConVibNet 的参数设置： $\alpha = 0.5$ 和 $\beta = 0.02$。Err.：误差；Suc.：成功。

未能检测到更深处的针尖，且低估了针长。相比之下，ConVibNet 在这两种情况下均能保持稳健的分割效果和更准确的针尖定位，即使在具有挑战性的可见性场景中也是如此。

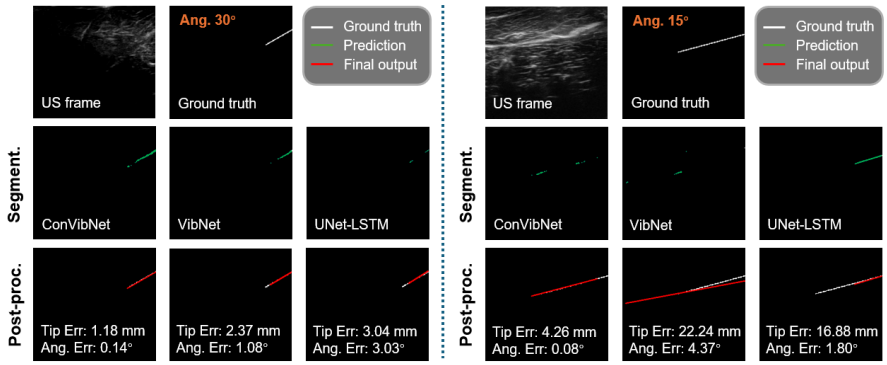


图 4：预测分割结果与后处理输出的示例比较

3.2 消融研究

为评估式（2）中所提出的各项损失函数的独立贡献，我们开展了系统的消融实验，结果汇总于表 3。除非另有说明，所有消融实验均采用相同的网络架构、训练计划、数据划分和数据增强策略，仅改变损失配置。作为基准的参考框架（“不使用 $\mathcal{L}_{\text{inter}}$ 和 $\mathcal{L}_{\text{diff}}$”）在不采用交集与差异损失项的情况下进行训练，其尖端误差为 3.55 ± 2.65 mm，角度误差为 $1.59 \pm 2.13^{\circ}$，成功率为 63。7%。

我们首先考察由 α [控制的交集损失对性能的影响。该损失（参见式（2））],旨在通过确保重叠输入序列之间的一致性，提升细粒度区域的准确性（参见“使用 $\mathcal{L}_{\text{inter}}$，不使用 $\mathcal{L}_{\text{diff}}$”）。引入交集损失显著提升了性能，并表明其起主导作用。在 $\alpha = 0.50$ 时取得了最佳结果，尖端误差显著降至 2.93 ± 2.54 毫米，成功率则升至 81。4%。然而，

表 3：交集与差异损失的消融研究

ConVibNet 框架	α	β	尖端误差 (mm)	角度误差 (°)	成功率 (%)
不含 $\mathcal{L}_{\text{inter}}$ 和 $\mathcal{L}_{\text{diff}}$	0.00	0.00	3.55 ± 2.65	1.59 ± 2.13	63.7
含 $\mathcal{L}_{\text{inter}}$，不含 $\mathcal{L}_{\text{diff}}$	0.25	0.00	3.66 ± 2.85	1.87 ± 2.34	66.9
	0.50	0.00	2.93 ± 2.54	1.78 ± 2.24	**81.4**
	0.75	0.00	3.49 ± 2.90	1.86 ± 2.24	77.0
含 $\mathcal{L}_{\text{inter}}$ 和 $\mathcal{L}_{\text{diff}}$	0.50	0.01	2.83 ± 2.45	1.76 ± 2.12	80.5
	0.50	**0.02**	$\mathbf{2.80 \pm 2.42}$	$\mathbf{1.69 \pm 2.00}$	79.6
	0.50	0.04	2.94 ± 2.86	1.73 ± 2.21	69.2
不含 $\mathcal{L}_{\text{inter}}$，含 $\mathcal{L}_{\text{diff}}$	0.00	0.02	4.04 ± 2.71	2.01 ± 2.51	58.3

角度误差和尖端误差以均值±标准差的形式描述；Err.：误差；Suc.：成功；w/o：不含；w/：含。

较低的权重（$\alpha = 0.25$）或较高的权重（$\alpha = 0.75$）都会导致性能变差，凸显了仔细调参的必要性。

将 α 固定在其最优值 0。50 后，我们通过系统地改变 β [（见公式 (2)], ，分析由其控制的差异损失对整体性能的影响。最佳整体性能由 β的组合实现，该组合为 α =0。50 和 $\beta = 0.02$，其尖端误差最低，为 2.80 ± 2.42 mm，角度误差为 $1.69 \pm 2.00°$。这表明，差异损失是一种有效的正则化项，可促使模型捕捉时间动态。关键的是，单独使用差异损失时（见“不含 $\mathcal{L}_{\text{inter}}$，含 $\mathcal{L}_{\text{diff}}$”），性能显著下降，尖端误差增至 4.04 ± 2.71 mm。这说明差异损失并非独立起作用，而是一种利用空间一致特征的有效正则化项。

总之，消融实验表明，两个损失组成部分都发挥着重要且互补的作用。交集损失主要通过保持空间一致性来提升性能，而差异损失则作为时间正则项，逐步优化结果。使用α=0.5 和 β=0.02 时性能最佳，在准确性与鲁棒性之间取得了良好平衡。这些发现证实了所提出的交集与差分损失设计的有效性。

4 讨论与结论

所提出的 ConVibNet 在典型临床超声帧率下实现了连续针检测的实时性能。除了具备计算效率外，ConVibNet 还实现了出色的针尖定位精度（$2.80 \pm 2.42\ mm$），成功率达到 79.6%；与表现最佳的基线方法相比，针尖精度提高了 0.75 mm，成功率提升了 15.9%。这些结果证实，所提出的交集-差异损失能够增强对针体运动的时间感知能力，从而在具有挑战性的成像条件下保持稳健的性能。

尽管取得了这些令人鼓舞的成果，仍存在若干局限。首先，数据收集受到高质量标注难以获取的限制，即使借助 NDI跟踪系统也是如此；此外，当前设置可用的插入角度范围有限（15°和30°）。选择这些角度是为了确保轨迹可重复，而非穷尽所有临床上可能的插入角度。因此，无法充分评估针杆角度预测的性能。此外，本研究未明确处理针体弯曲和操作者差异的影响。研究也未验证模型在其他探头类型、超声系统及针规格上的泛化能力。因此，未来研究将探究更多样化的插入设置，将针体弯曲纳入检测框架，评估人为引起的差异所产生的影响，并验证模型在不同成像配置和针规格下的泛化能力。构建配备光纤传感器的机器人平台，有望在插入过程中同时捕捉针姿态和变形，从而进一步促进可靠真实值的获取。另一项方法学局限在于模型对运动动力学的隐式处理。观测到的针体运动有两个来源：操作者操控，以及叠加用于信号增强的高频振动。由于当前模型未能区分这两种成分，其学习完全稳健表征的能力可能受到限制。未来能够明确分离这些运动来源的架构，或可实现更准确且泛化能力更强的预测。

总之，我们提出了 ConVibNet，这是针对超声引导介入中的连续针检测而扩展的 VibNet。通过新颖的交集-差异损失设计和时间相关性建模，ConVibNet 在保持实时性能的同时，显著提升了准确性和鲁棒性。这些进展凸显了其作为可集成到电动进针系统中的可行候选方案的潜力，也是迈向自主超声引导手术的有力一步。

声明

利益冲突

作者声明不存在任何利益冲突。

资金支持

本研究由多尺度医疗机器人中心、AIR@InnoHK（香港）资助。

数据可用性

支持本研究结论的数据可根据合理请求向通讯作者索取。

参考文献

[1] Kuang, Y., Hilgers, A., Sadiq, M., Cochran, S., Corner, G., Huang, Z.: 对一种超声驱动针进行建模

并表征其特性，以提升其在超声引导下区域麻醉和组织活检中的可视性。《Ultrasonics》**69**，38–46（2016）[2] Huang, D., Li, C., Karlas, A., Chu, X., Samuel Au, K.W., Navab, N., Jiang, Z.: Vibnet：利用振动增强超声图像中的针检测。《IEEE医学影像学汇刊》**44**(6)，2696–2708（2025）[3] Zhang, Y., Lei, L., Yan, W., Zhang, T., Tang, R.S.-Y., Cheng, S.S.: Mambax-ctrack：结合 SSM 交叉相关和运动提示的基于 Mamba 的跟踪器，用于超声针跟踪。《IEEE机器人与自动化快报》（2025）[4]Goel, R., Morales, C., Singh, M., Dubrawski, A., Galeotti, J., Choset, H.: 超声图像中基于运动信息的针分割。arXiv 预印本 arXiv:2312.01239（2023）[5] Mathiassen, K., Dall'Alba, D., Muradore, R., Fiorini, P., Elle, O.J.: 使用统计滤波在二维超声图像中实现稳健的实时针跟踪。《IEEE控制系统技术汇刊》**25**(3)，966–978（2016）[6] Kaya, M., Senel, E., Ahmad, A., Orhan, O., Bebek, O.: 面向机器人活检的二维超声图像实时针尖定位。载于：2015年先进机器人国际会议（ICAR），第47–52页（2015）。IEEE[7] Jiang, Z., Salcudean, S.E., Navab, N.: 机器人超声成像：现状与未来展望。《医学图像分析》**89**，102878（2023）[8] Bi, Y., Jiang, Z., Duelmer, F., Huang, D., Navab, N.: 机器人超声成像中的机器学习：挑战与展望。《控制、机器人与自主系统年度评论》**7**（2024）[9] Jiang, Z., Bi, Y., Zhou, M., Hu, Y., Burke, M., Navab, N.: 智能机器人超声医师：基于互信息的解耦奖励学习，从少量示范中学习。《国际机器人研究杂志》**43**(7)，981–1002（2024）[10] Mwikirize, C., Nosher, J.L., Hacihaliloglu, I.: 用于二维超声实时针检测与定位的卷积神经网络。《国际计算机辅助放射学与外科学杂志》**13**(5)，647–657（2018）[11] Wijata, A., Andrzejewski, J., Pyci´nski, B.: 利用卷积神经网络在超声图像中自动检测与分割活检针。《超声成像》**43**(5)，262–272（2021）[12] Huang, D., Bi, Y., Navab, N., Jiang, Z.: 机器人超声检查中的运动放大：实现脉动感知动脉分割。载于：2023年IEEE/RSJ智能机器人与系统国际会议（IROS），第6565–6570页

(2023)。IEEE

[13]Belikova, K., Zailer, A., Tekucheva, S.V., Ermoljev, S.N., Dylov, D.V.：利用深度学习在超声视频中对颞下颌关节进行时空定位。载于：2021 IEEE生物信息学与生物医学国际会议（BIBM），第1257–1261页（2021）。IEEE[14]Hovgesen, C.H., Wilhjelm, J.E., Vilmann, P., Kalaitzakis, E.：用于改善超声针可视化的增强回声表面：《超声》PRISMA系统综述。《医学超声杂志》**41**(2), 311–325 (2022)[15] Dupere, J.M., Brost, E.E., Uthamaraj, S., Lee, C.U., Urban, M.W., Stish, B.J., Deufel, C.L.：利用超声彩色多普勒和针表面改性可视化前列腺近距离放射治疗针的新方法。《近距离放射治疗》**22**(6), 761–768 (2023)[16] Beigi, P., Rohling, R., Salcudean, S.E., Ng, G.C.：通过时空线性采样对震颤运动进行谱分析，以在曲线阵超声中检测针。《国际计算机辅助放射学与外科学杂志》**11**(6), 1183–1192 (2016) [17] Beigi, P., Rohling, R., Salcudean, T., Lessoway, V.A., Ng, G.C.：利用概率支持向量机和时域特征检测超声中的不可见针。《Ultrasonics》**78**, 18–22 (2017)[18] Lin, T.-Y., Goyal, P., Girshick, R., He, K., Doll´ar, P.：用于密集目标检测的焦点损失。载于：2017 IEEE计算机视觉国际会议（ICCV），第2999–3007页（2017）[19] Jiang, Z., Li, X., Chu, X., Karlas, A., Bi, Y., Cheng, Y., Au, K.S., Navab, N.：利用生成对抗网络进行针分割：恢复机器人超声中的细长器械可视化。《IEEE仪器与测量汇刊》（2024）[20] Huang, D., Yang, C., Zhou, M., Karlas, A., Navab, N., Jiang, Z.：利用虚拟夹具进行机器人辅助深静脉血栓超声检查。《IEEE自动化科学与工程汇刊》**22**, 381–392 (2024)[21] Abolhassani, N., Patel, R., Moallem, M.：软组织针插入：综述。《医学工程与物理》**29**(4), 413–431 (2007)