

基于智能手机CSI的轻量级指纹定位模型

杨润泽^{1, 2, 3}, 黄宝琦^{1, 2, 3}, 贾冰^{1, 2}

(1. 内蒙古大学计算机学院 呼和浩特 010021; 2. 生态大数据教育部工程研究中心 呼和浩特 010021; 3. 内蒙古电子科技有限公司 鄂尔多斯 017000)

摘要: 近年来, 深度学习技术被广泛地应用于基于信道状态信息(Channel State Information, CSI)的指纹定位领域, 并展现出高可靠的定位精度。然而, 基于深度学习的指纹定位方法大多依赖于多层次的网络模型结构来提取具有判别性的位置特征, 此过程伴随着大量的模型参数和密集的计算操作, 对设备的硬件资源占用较高。特别是对于资源有限的智能手机而言, 这会为其带来不容忽视的负担。为此, 本文提出一种新颖的智能手机CSI轻量级指纹定位模型, (1) 通过设计简洁高效的特征提取模块, 仅用少量参数的线性层与卷积层结合特征融合机制, 兼顾模型轻量化与特征提取能力; (2) 引入数据增强模块, 通过生成非参考点CSI数据, 扩充训练样本空间, 显著提升模型对未经训练位置的鉴别能力。本文在两种典型的室内场景下开展实验, 与最优的基准模型相比, 本文提出的定位模型在两种场景下的平均RMSE降低了12.5%, 并且单次定位的推理时间仅为0.08秒。实验结果表明, 本文提出的定位模型不仅能够有效提高定位精度, 还明显缩短了定位所需时间。

关键词: 智能手机CSI; 室内定位; 深度学习; 轻量级模型

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.2096-3750.XXXX.

A Lightweight Fingerprinting Localization Model based on Smartphone CSI

YANG Runze^{1,2,3}, HUANG Baoqi^{1,2,3}, JIA Bing^{1,2}

1. College of Computer Science, Inner Mongolia University, Hohhot, 010021, China

2. Engineering Research Center of Ecological Big Data, Ministry of Education, Hohhot, 010021, China

3. Inner Mongolia Electronic Technology Co., LTD, Ordos, 017000, China

Abstract: In recent years, deep learning technologies have been widely adopted to the Channel State Information (CSI)-based fingerprinting localization field, demonstrating reliable localization accuracy. However, deep learning-based fingerprinting localization models typically rely on multi-layered network structures to extract distinctive location features, the process accompanied by a large number of model parameters and intensive computational operations, which occupy substantial physical resources on devices. This can be a significant burden, especially for resource-limited smartphones. To this end, this paper proposes a novel lightweight fingerprinting localization model based on smartphone CSI, (1) a concise and efficient feature extraction module is designed, employing linear and convolutional layers with a minimal number of parameters and incorporating a feature fusion mechanism to achieve both lightweight deployment and strong discriminative capability in feature extraction; (2) a data augmentation module is introduced to generate CSI data for unknown

收稿日期: 2025-06-XX; 修回日期: 2025-XX-XX

通信作者: 黄宝琦, cshbq@imu.edu.cn

基金项目: 国家自然科学基金项目(No.62262046, No.42161070); 内蒙古科技计划项目(No.2022YFSJ0027); 鄂尔多斯市科技计划项目(No.YF20240029); 呼和浩特市基础研究与应用基础研究项目(No.2025-规-基-42)

Foundation Items: The National Natural Science Foundation of China (No. 62262046, No. 42161070), Inner Mongolia Science and Technology Program (No. 2022YFSJ0027), Ordos Science and Technology Program (No. YF20240029), and Hohhot Basic and Applied Basic Research Program (No. 2025-GUI-JI-42)

points, thereby expanding the training sample space and significantly enhancing the model's recognition ability for untrained locations. Experiments were conducted in two typical indoor scenarios. The results show that compared to the most advanced benchmark models, the proposed localization model has reduced the average RMSE by 12.5% in two indoor scenarios, and the inference time for a single localization is only 0.08 seconds. The proposed localization model not only effectively improves positioning accuracy but also significantly reduces the required positioning time.

Key words: smartphone channel state information, indoor localization, deep learning, lightweight model

0 引言

无线室内定位技术利用 WiFi、RFID、UWB 和蓝牙等射频信号来估算待定位目标在室内环境中的真实位置^[1]。其中, WiFi 定位方法具有普适性高和低成本等显著优势。例如, 购物中心或办公区域通常安装大量的可用于提供位置服务的 WiFi 接入点 (Access Point, AP), 从而降低此类硬件的部署成本。目前, WiFi 室内定位技术被广泛应用于室内跟踪^[2]、健康监测^[3]和虚拟现实^[4]等领域。

早期的 WiFi 定位方法主要通过接收信号强度 (Received Signal Strength, RSS) 来估计目标位置。然而, RSS 仅能粗略地反映信号强弱, 并且容易受到多径效应的干扰, 这些因素共同限制了此方法的准确性和鲁棒性。在采用正交频分复用的 WiFi 系统中, 信道状态信息 (Channel State Information, CSI) 能够提供单个信道中不同子载波的详细信息。由于这些子载波相互正交, CSI 具有更强的抗干扰能力, 从而能够应对多径效应的影响^[5]。此外, CSI 还详细地反映了无线信号在传播过程中的相位和幅度变化, 为实现高精度定位提供了丰富信息^[6]。FILA^[7]是最早比较 CSI 和 RSS 定位性能的工作, 它通过修正的传播模型和三边定位法进行定位。实验结果表明, CSI 定位的准确性明显优于 RSS 定位。综上所述, CSI 定位方法在定位性能上具有显著优势, 并在室内定位领域展现出了广阔的应用前景。

目前, 基于 CSI 的室内定位方法主要分为几何测距法和指纹定位法。前者的核心思想是通过计算目标与信号收发端之间的到达角 (Angle of Arrival, AoA)、离开角 (Angle of Departure, AoD) 和飞行时间 (Time of Flight, ToF) 来推算位置。例如, SpotFi^[8]采用超分辨率算法准确计算多径信号的 AoA 和 ToF, 并成功识别出在视距 (Line-of-Sight, LoS) 路径下目标与 AP 之间的 AoA, 以实现亚米级的定位精度。MonoLoco 系统^[9]提出三维超分辨率

算法估计多径信号的 AoA、AoD 和 ToF 等特征, 构建一个由接收器、目标设备及反射器共同组成的几何图形, 通过解析几何图形关系, 实现对目标设备的准确定位。mD-Track^[10]系统将 AoA、ToF 和多普勒频移等信息进行融合, 在不需要更宽的频率带宽和更多天线数量的情况下实现了高精度定位。尽管几何距离法能够提供较好的定位精度, 但该方法在实际应用中可能受到非视距 (None Line-of-Sight, NLoS) 因素 (如人员移动和物理遮挡) 的影响^[11]。相比之下, 指纹定位法无需预知 AP 的确切位置, 也不用进行复杂的时间或角度测量, 展现出了较好的灵活性和实用性。因此, 该方法受到了研究人员的广泛关注。

传统的指纹定位法通常在离线阶段将每个参考点 (Reference Point, RP) 处的 CSI 测量的统计值 (如平均值、中位数等) 作为指纹, 并将所有 RP 上的 CSI 指纹整合以构建与真实位置空间相对应的指纹库。在定位阶段, 利用 KNN (K-Nearest Neighbor)^[12] 或朴素贝叶斯 (Naive Bayes, NB)^[13-14] 等机器学习 (Machine Learning, ML) 技术将接收到的 CSI 测量与指纹库中的 CSI 指纹进行匹配, 以估计位置。Pilot^[14] 是首个基于 CSI 的指纹定位系统, 它在离线阶段利用 CSI 幅度因环境波动所引起的信号变化, 构建位置空间中的指纹。在定位阶段, Pilot 采用贝叶斯估计方法来推算目标所处的位置。然而, 由于 CSI 测量的维度较高, 其中有价值的数据可能受到环境噪声干扰, 使得 ML 技术难以提取出更具代表性的位置特征, 从而限制了此类方法的定位精度。

随着人工智能领域的快速发展, 深度学习 (Deep Learning, DL) 技术能够自动检测和提取具有强大表征能力的特征。因此, 研究人员开始逐渐采用更先进的 DL 技术取代传统的 ML 技术。例如, 卷积神经网络 (Convolutional Neural Network, CNN)^[15, 16] 能够用于 CSI 指纹定位领域, 它通过捕捉 CSI 信号在不同位置空间上的变化, 从而实现准确

定位。此外，同一位置上的CSI测量值会随着时间的推移而发生变化。由于长短期记忆网络(Long Short-Term Memory, LSTM)^[17]能够捕捉时间序列数据中存在的长期依赖关系，因此LSTM同样被应用于CSI指纹定位领域。

目前，多数研究^[15, 16, 17]通常利用配备无线网卡(Intel 5300或Atheros 9462)的台式机或笔记本电脑采集CSI测量，存在一定的局限性。首先，适用于专用无线网卡的CSI工具通常仅支持IEEE 802.11n标准，其对应的最大带宽为40 MHz。相比之下，适用于智能手机的CSI工具除了支持IEEE 802.11n标准外，还支持更高效的802.11ac标准，使其能够在高达80 MHz的带宽下解析出包含更多子载波的CSI测量，它们蕴含着更高分辨率的信息。利用这些信息有助于提升模型的位置分辨能力，使其能够更准确地识别和区分微小的空间位置变化^[18]。其次，与智能手机相比，台式机等设备较为笨重，不便于携带和移动。鉴于智能手机的便携性和普适性，尤其是对于每天频繁使用智能手机的用户来说，采用智能手机CSI获取室内位置信息成为理想选择。然而，由于基于DL的指纹定位模型^[15, 16, 17]通常具有复杂的多层网络结构，因此它们在处理高维CSI测量的过程中需要占用大量的计算资源，这对于资源有限的智能手机而言是一个重大挑战。

为此，本文提出基于智能手机CSI的轻量级指纹定位模型LFLM-SCSI(Lightweight Fingerprinting Localization Model based on Smartphone CSI)，包含特征提取和数据增强两个关键模块。首先，特征提取模块旨在通过简洁的网络结构设计，提取CSI测量中与位置相关的判别性特征，以尽量避免对复杂模型结构的依赖。为此，特征提取模块采用基于直接多步预测的线性层网络NLinear和典型的残差块(Residual Block, RB)结构，以分别捕获CSI测量的整体趋势特征和局部空间特征。同时，设计一种特征融合机制，在不增加模型参数数量的前提下高效地融合上述特征。由于特征提取模块的简洁网络结构可能会限制模型的定位精度，因此提出数据增强模块生成扩展点(Extended Point, EP)上的CSI测量，从而为特征提取模块提供定位区域内更为全面的位置信息，改善定位结果的鲁棒性。此外，本文还引入树突(Dendritic, DD)网络，以捕获位置空间与输入特征之间的逻辑关系。本文的主要贡献总结

如下：

(1) 首次提出基于智能手机CSI的轻量级定位模型。由于该模型在定位过程中主要依赖于包含可并行处理的线性卷积操作的特征提取模块，因此与其它网络结构复杂的DL模型相比，该模型在定位过程中能明显减少其对资源的需求。

(2) 数据增强模块结合高斯过程回归(Gaussian Process Regression, GPR)方法和条件变分自编码器(Conditional Variational Autoencoder, CVAE)模型，通过使用GPR估算EP位置处的CSI指纹，并将这些指纹作为条件变量和随机变量一起输入到生成模型的解码器中，从而在EP位置处生成可靠且多样化的CSI测量。

(3) 在两种典型的室内场景下进行实验。结果表明，与定位精度最好的基线模型相比，LFLM-SCSI在两种室内场景下的平均RMSE降低12.5%，并且完成单次定位的时间约为0.08秒。

1 相关工作

1.1 基于ML的指纹定位方法

目前，多种ML技术以监督学习方式被广泛应用于CSI指纹定位领域，包括但不限于KNN、支持向量机(Support Vector Machine, SVM)、NB和随机森林(Random Forest, RF)。

KNN和SVM。Song等人^[19]通过多维尺度分析(Multi-Dimensional Scaling, MDS)来计算用户与RP之间的欧氏距离和时间反转共振强度，进而将MDS与KNN相结合以估计位置。FuseLoc^[20]利用典型相关分析在特征层面融合由CSI幅度和相位提取的位置特征，并通过KNN进行位置估计。Zhou等人^[21]首先采用SVM建立CSI测量与其对应位置间的非线性关系，然后通过聚类方法以减少CSI测量中噪声的影响，并使用主成分分析(Principal Component Analysis, PCA)来降低特征维度。

NB和RF。FapFi^[22]方法采用一种简单的数据融合策略，该方法直接在数据层面融合CSI的幅度和相位测量，然后利用NB算法进行位置预测，实现相对较低的定位误差。另一项值得注意的工作是RFFP^[22]，它通过在离线阶段训练RF模型，并将此模型用于存储位置指纹。这样做不仅有助于节省算法所需的存储空间，还能够提升此方法应对多径效应的能力。

尽管 ML 技术在提升定位精度方面展现出了一定优势,但这些方法往往依赖手工提取的统计特征,忽略了原始 CSI 测量中隐藏的关键信息,而这些信息对于定位任务而言至关重要。

1.2 基于 DL 的指纹定位方法

随着人工智能领域的快速发展,DL 技术能够自动从高维 CSI 测量中提取有意义的位置特征,从而逐渐凸显其在指纹定位领域中的重要性。Wang 等人^[23-25]将 DL 技术应用于 CSI 指纹定位领域,并取得了一系列研究成果。PhaseFi^[23]设计了一个带有三层隐藏层的深度网络,用于训练校准后的 CSI 相位,取得了比 FILA 更好的定位精度。DeepFi^[24]采用多层感知器 (Multi-Layer Perceptron, MLP) 将网络权重作为 CSI 指纹,并通过贪婪学习算法进行权重训练。类似地, BiLoc^[25]融合了 CSI 幅度的平均测量以及由 CSI 相位估计出的几何信息,通过深度自编码器网络处理融合数据,并将网络中的权重作为每个位置的指纹。

近年来, CNN、LSTM 和它们的变体模型在 CSI 指纹定位领域得到了广泛应用。例如, CNN-Loc 和 DCNN^[15-16]通过多层卷积结构强化位置特征的提取能力。Choi 等人^[17]采用 LSTM 模型,根据不同比例的 CSI 幅度和相位测量来预测位置。Hoang 等人^[26]将 CNN 模型与 LSTM 模型相结合,从智能手机采集的 CSI 幅度图中提取时空特征,利用 CNN 的空间特征提取能力和 LSTM 的时间序列处理能力,提升定位精度。Liu 等人^[27]提出了 FTLoc 定位模型,利用时序特征提取模块、注意力泛化机制和特征融合模块,改善定位性能。随着大模型和第六代移动通信技术的出现,CSI 指纹定位技术也随之进一步发展。近期的指纹定位研究采用残差网络和 Transformer 模型处理高维 CSI 测量,实现分米级定位精度。Fan 等人^[28]提出了室内定位模型 SBi-LSTM (Structured Bidirectional LSTM),该网络结合残差网络和 Bi-LSTM 结构,有效提升了 NLOS 情况下的定位精度和鲁棒性。SwinLoc^[29]采用了先进的 Swin-Transformer^[30]架构,融合了自注意力机制和多尺度卷积,能够自适应地捕捉 CSI 测量中的细微多尺度模式,以降低定位误差。

综上所述,近期的 CSI 指纹定位研究通过 DL 技术使得定位精度取得了明显提升。然而,采用结

构复杂的深度学习模型处理高维的 CSI 测量势必会导致大量计算资源消耗^[31],这对于资源有限的智能手机而言是一项重大挑战。因此,本文旨在采用智能手机 CSI 开发一种轻量级定位模型,在保持定位精度的同时减少对移动设备的资源需求。

2 预备知识

2.1 智能手机 CSI

为使用智能手机采集 CSI 测量,需要在手机上安装 CSI 提取工具,即 Nexmon CSI^①。此工具支持内置 Broadcom WiFi 芯片的智能手机(如 Nexus 6P)从监听的 WiFi 帧中提取相应的 CSI 测量。利用 Nexmon CSI 从 WiFi 帧中解析出来的 CSI 测量可以表示为 $\mathbf{H} = [h_1, \dots, h_k, \dots, h_K]$, 其中 $h_k = a_k e^{j\phi_k}$ 。其中, a_k 和 ϕ_k 表示第 k 个子载波上的幅度和相位,它们分别反映信道衰减和相位偏移。理论上,幅度和相位可用于推断信号的传播距离和方向。然而,原始的 CSI 相位测量由于受到采样频偏和载波频偏的影响而失真,使其不能直接用于指纹定位^[32]。此外,由于数据子载波上的 CSI 幅度测量不受异常值的干扰¹,因此本文仅采用数据子载波上的 CSI 幅度测量,即 $\mathbf{A} = [a_1, \dots, a_{K'}]$ ($K' < K$), 用于位置估计。值得注意的是,智能手机 CSI 测量与无线网卡采集的 CSI 测量之间的数据格式存在差异。智能手机 CSI 测量的数据格式为 $\mathbf{H} \in \mathbb{C}^{K \times N_w}$, 其中 K 和 N_w 分别代表子载波数量和 CSI 样本数量,而该格式并不包含发送端和接收端的天线数量。

2.2 基于 DL 的 CSI 定位

基于 DL 的 CSI 定位原理与传统的指纹定位方法类似,同样包含两个阶段。在离线阶段,需要在每个 RP 处采集 N_w 个 CSI 幅度测量 $\{\mathbf{A}_1, \dots, \mathbf{A}_{N_r}\}$ 用于训练定位模型。假设 RP 的数量为 N_r , 则采集到的用于训练定位模型的 CSI 数据集可以表示为 $\mathcal{Q} \in \mathbb{R}^{M \times K'}$, 其中 $M = N_r \times N_w$ 。然后,将 N_r 个 RP 的位置坐标 $\mathcal{P} = \{\mathbf{L}_1, \dots, \mathbf{L}_{N_r}\}$ 与 \mathcal{Q} 结合构成训练数据集,其中 \mathbf{L} 表示二维位置坐标。接下来,根据数据集训练定位模型,建立 CSI 幅度测量与其对应的位置坐标之间的映射关系,即 $f(\mathbf{A}; \theta) = \tilde{\mathbf{L}}$, 其中 θ 表示模型参数, $\tilde{\mathbf{L}}$ 表示预测位置。为了获得准确的定位结果,需要最小化真实位置 \mathbf{L} 与估计位置 $\tilde{\mathbf{L}}$ 之

①Nexmon CSI. https://github.com/seemoo-lab/nexmon_csi.

间的定位误差。该误差通常由均方根误差(Mean Squared Error, MSE)来衡量, 即:

$$L_{MSE} = \frac{1}{M} \sum_{m=1}^M |L_m - \tilde{L}_m|^2 \quad (1)$$

在此过程中, 模型通过梯度下降或其他优化算法调整网络参数 θ , 使得距离相近的位置特征尽可能接近, 而距离较远的位置特征尽可能远离。在定位阶段, 将接收到的CSI幅度测量输入到训练好的定位模型中, 即可得到相应的估计位置 \tilde{L} 。

然而, 基于DL的指纹定位方法通常更聚焦于定位精度, 对深度网络的结构复杂性关注较少。为了从高维CSI测量中提取有意义的位置特征, 基于DL的CSI指纹定位研究^[15-17,27-29]通过级联具有大量参数和复杂计算的卷积网络、残差网络和Transformer模型来构造多层特征提取模块, 但它们却难以在资源受限的设备上高效部署。

3 轻量级定位模型

3.1 模型概述

为减轻定位模型对硬件资源的依赖, 直观的思路是设计一个相对简单的深度网络结构, 以降低模型对硬件资源的占用。因此, 本文提出了一种轻量级定位模型LFLM-SCSI, 采用简单的线性层和卷积层结构搭建轻量化的位置特征提取模块, 通过少量参数和简单的运算快速提供位置信息。然而, 简单网络的特征学习能力有限, 可能导致定位精度不足, 难以满足实际应用需求。为此, 本文还设计了数据增强模块, 采用此模块在模型训练阶段提供更加丰富和多样化的位置特征, 以增强特征提取模块的学习能力。

具体而言, LFLM-SCSI定位模型包括数据增强模块、特征提取模块和位置估计模块, 模型整体结构如图1所示。首先, 智能手机CSI幅度测量 \mathbf{A} 被输入到数据增强模块, 通过将RP处的CSI幅度测量输入到GPR模型中, 以估计指定EP处的CSI指纹 $\bar{\mathbf{A}}^*$, 并将其作为条件变量输入到CVAE生成模型以仿真生成EP处的CSI幅度测量 \mathbf{A}^* 。其次, 将RP处真实测量 \mathbf{A} 与EP处生成测量 \mathbf{A}^* 合并输入到特征提取模块, 以进一步提升模型对定位空间的分辨能力。换言之, 通过有规律地选取定位空间中的EP位置, 并生成其位置上的CSI幅度测量, 增强了模型对未知位置的识别能力, 同时减轻了数据采集

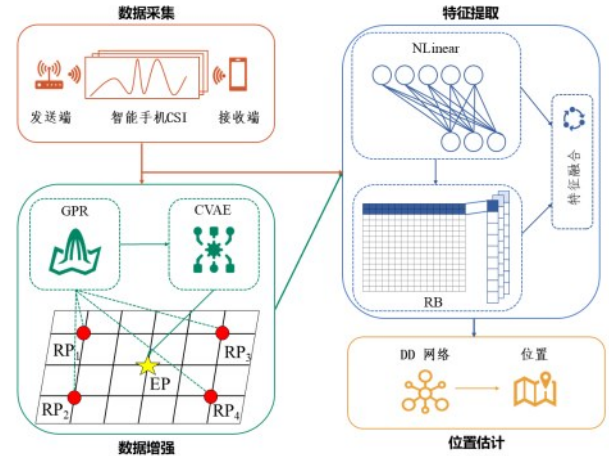


图1 LFLM-SCSI的整体模型结构

的人力成本。特征提取模块采用NLinear层和RB结构来捕获CSI幅度测量与其对应位置之间的整体趋势特征和局部相关性, 并通过本文提出的特征融合机制进行处理。最后, 通过DD网络对融合后的特征进行特征交互, 从而获取融合特征与位置标签之间的关系映射, 以得到准确的定位结果。

此外, 在将CSI数据集 \mathcal{Q} 输入到数据增强模块之前, 本文首先使用移动中值滤波器对CSI幅度测量在时间维度上的尖锐噪声进行处理, 然后对滤波后的数据使用最大最小归一化方法来提高模型的收敛速度和整体性能。

3.2 数据增强

3.2.1 估计CSI指纹

数据增强模块主要采用CVAE的思想^[33]以尽可能准确地生成EP处的CSI幅度测量。在此过程中, 需要将EP处的CSI指纹作为条件变量, 以确保在EP处生成的CSI幅度不仅在统计特性上与RP的训练数据相似, 而且能够客观地反映其在相应的位置空间上的变化趋势。然而, 由于EP处的CSI指纹是未知的, 因此本文采用经典的GPR方法^[34], 结合RP上的CSI幅度测量估计合适的模型参数, 以此估计EP处的CSI指纹。

假设第 i 个RP处的CSI指纹 $\bar{\mathbf{A}}_i = \sum_{n=1}^{N_s} \mathbf{A}_n$, 并且 $\bar{\mathbf{A}}_i$ 与其位置标签 L_i 之间的映射关系为 $\bar{\mathbf{A}}_i = g(L_i)$, 那么位置标签 L 及其对应的CSI指纹 $\bar{\mathbf{A}}$ 的高斯过程可以表示为:

$$g(L) \sim GP(\mu(L), \kappa(L, L_j)) \quad (2)$$

其中, $\mu(\cdot)$ 和 $\kappa(\cdot)$ 分别表示均值函数和协方差

函数，即 $\mu(\mathbf{L}) = E[g(\mathbf{L})]$ ， $\kappa(\mathbf{L}_i, \mathbf{L}_j) = E[(g(\mathbf{L}_i) - \mu(\mathbf{L}_i))(g(\mathbf{L}_j) - \mu(\mathbf{L}_j))]$ 。假设观测误差 ϵ 是服从均值为零且方差为 σ_r^2 的高斯白噪声，即 $\epsilon \sim N(0, \sigma_r^2)$ ，则位置标签 \mathbf{L} 与其对应的 CSI 指纹 $\bar{\mathbf{A}}$ 之间的关系可以表示为 $\bar{\mathbf{A}} = g(\mathbf{L}) + \epsilon$ 。

为了简化后验概率分布的求解，GPR 通常将均值函数设置为零。为此，本文采用 Z-Score 方法对 RP 处的 CSI 指纹进行去中心化，因此 $\bar{\mathbf{A}}$ 的先验分布为： $\bar{\mathbf{A}} \sim N(0, \kappa(\mathbf{L}_i, \mathbf{L}_j) + \sigma_r^2)$ 。根据 GPR 原理， $\bar{\mathbf{A}}$ 和 EP 处的 CSI 指纹 $\bar{\mathbf{A}}^*$ 存在相关性，并服从联合高斯分布：

$$\begin{bmatrix} \bar{\mathbf{A}} \\ \bar{\mathbf{A}}^* \end{bmatrix} \sim N \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \kappa(\mathbf{L}, \mathbf{L}) + \sigma_r^2 & \kappa(\mathbf{L}, \mathbf{L}^*) \\ \kappa(\mathbf{L}^*, \mathbf{L}) & \kappa(\mathbf{L}^*, \mathbf{L}^*) \end{bmatrix} \right\}, \quad (3)$$

其中， $\kappa(\mathbf{L}, \mathbf{L}^*) = \kappa(\mathbf{L}^*, \mathbf{L})$ 代表 RP 与 EP 之间的协方差矩阵； $\kappa(\mathbf{L}^*, \mathbf{L}^*)$ 代表 EP 的协方差矩阵。根据联合高斯分布条件分布的相关结论^[35]，可知 EP 处的 CSI 幅度测量 $\bar{\mathbf{A}}^*$ 的后验分布为：

$$P(\bar{\mathbf{A}}^* | \mathbf{L}, \bar{\mathbf{A}}, \mathbf{L}^*) \sim N(\mu^*, (\sigma^*)^2), \quad (4)$$

其中 $\mu^* = \kappa(\mathbf{L}^*, \mathbf{L}) [\kappa(\mathbf{L}, \mathbf{L}) + \sigma_r^2]^{-1} \bar{\mathbf{A}}$ ， $(\sigma^*)^2 = \kappa(\mathbf{L}^*, \mathbf{L}^*) - \kappa(\mathbf{L}^*, \mathbf{L}) [\kappa(\mathbf{L}, \mathbf{L}) + \sigma_r^2]^{-1} \kappa(\mathbf{L}, \mathbf{L}^*)$ 。因此，EP 处的 CSI 指纹可以通过后验均值 μ^* 和方差 $(\sigma^*)^2$ 计算得到。

通常，GPR 利用核函数构造高斯过程中的协方差矩阵。常见的核函数为径向基函数(Radial Basis Function, RBF)：

$$\kappa_{\text{RBF}}(\mathbf{L}_i, \mathbf{L}_j) = \exp \left[-\frac{d(\mathbf{L}_i, \mathbf{L}_j)}{2\nu^2} \right], \quad (5)$$

其中 $d(\mathbf{L}_i, \mathbf{L}_j) = \|\mathbf{L}_i - \mathbf{L}_j\|^2$ ， ν 表示长度尺度参数。此外，除了 RBF 核，本文还考虑了 Matern 核和有理二次(Rational Quadrati, RQ)核以及它们的复合函数核，以灵活地捕捉 CSI 指纹的空间变化趋势。与 RBF 核相比，Matern 核有一个额外的参数 s ，用于控制输出结果的平滑度，其表达式为：

$$\kappa_{\text{Matern}}(\mathbf{L}_i, \mathbf{L}_j) = \frac{1}{\Gamma(s)2^{s-1}} \left(\frac{\sqrt{2s}}{\nu} d(\mathbf{L}_i, \mathbf{L}_j) \right)^s K_s \left(\frac{\sqrt{2s}}{\nu} d(\mathbf{L}_i, \mathbf{L}_j) \right), \quad (6)$$

其中 $\Gamma(\cdot)$ 表示伽玛函数， $K_s(\cdot)$ 表示修正贝塞尔函数。RQ 核的表达式为：

$$\kappa_{\text{RQ}}(\mathbf{L}_i, \mathbf{L}_j) = \left(1 + \frac{d(\mathbf{L}_i, \mathbf{L}_j)}{2\alpha\nu^2} \right)^{-\alpha}, \quad (7)$$

其中 α 是混合尺度参数。核函数及其参数选择对于准确捕捉 CSI 幅度测量中蕴含的环境变化特性至关重要。因此，本文通过交叉验证方法来确定最优的参数组合，以确保估计的位置指纹准确性。

值得注意的是，GPR 方法不能够直接用于未采集 RP 的新环境或环境剧烈变换场景。此外，该方法需要足够多的真实 CSI 指纹，才能实现对当前位置空间中未知位置 CSI 指纹的准确预测。若用于训练的 CSI 指纹不能够完全覆盖定位区域，则 GPR 难以准确地学习到 CSI 测量在位置空间中的变化趋势，从而影响生成的指纹质量。为此，本文设置的参考点最大程度上覆盖了定位区域，以确保生成指纹的准确性。

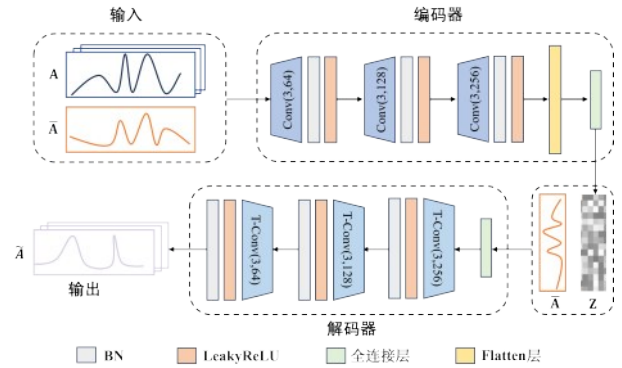


图2 CVAE生成模型结构

3.2.2 生成CSI测量

为了生成 EP 处的 CSI 幅度测量 \mathbf{A}^* ，本文首先利用 RP 处的 CSI 幅度测量及其指纹训练一个基于 CVAE 结构的生成模型，以得到能够准确模拟 CSI 幅度分布的生成器。一旦模型训练完成，便可通过向模型输入 EP 处的 CSI 指纹及随机变量来生成大量可用的 CSI 幅度测量。具体而言，在获得 RP 处的 CSI 指纹 $\bar{\mathbf{A}}$ 后，将其作为条件变量与 RP 处的 CSI 幅度 \mathbf{A} 一起输入到生成模型中，通过训练获得 CSI 幅度生成器。本文设计的生成模型整体结构包含一个编码器和一个解码器，如图 2 所示。模型训练完成后，训练好的解码器被作为产生 CSI 幅度测量的生成器。生成模型的具体训练步骤如下。

首先，编码器将 RP 处的 CSI 幅度 \mathbf{A} 及其 CSI 指

纹 \bar{A} 映射到潜在特征空间中,学习到潜在特征 z 的条件分布,此过程通过计算潜在变量的均值 μ 和方差 σ^2 来实现。为了通过反向传播优化模型,编码器利用重参数化技巧从标准正态分布中抽取随机噪声 η ,并通过公式 $z = \mu + \sigma \odot \eta$,得到潜在特征 z 。此过程不仅保持模型所需的样本随机性,同时也支持梯度的反向传播。

其次,将潜在特征 z 重采样并与条件变量 \bar{A} 拼接后一起输入到解码器中,从而生成与原始数据分布相似的CSI幅度测量 \tilde{A} 。为了尽可能地减少原始CSI幅度测量 A 和生成的CSI幅度测量 \tilde{A} 之间的差异,并确保重构的高维CSI测量的准确性。生成模型的整体优化目标为:

$$L_{CVAE} = -\text{KL}(q(z|A, \bar{A}) // P(z)) + E_{q(z|A, \bar{A})}(\lg P(A|z, \bar{A})), \quad (8)$$

其中, $\text{KL}(\cdot)$ 表示KL散度; $P(z)$ 是潜在变量 z 的先验分布,通常设置为标准正态分布; $q(z|A, \bar{A})$ 代表在给定输入 A 和 \bar{A} 条件下潜在变量 z 的条件概率分布; $E(\cdot)$ 表示期望; $P(A|z, \bar{A})$ 则是在给定潜在变量 z 和条件变量 \bar{A} 的情况下观测测量 A 的条件概率分布。公式(8)包含一个约束部分以及一个重构部分。前者用于提升内在特征 z 的多样性,而后者确保生成的CSI幅度测量 \tilde{A} 的准确性。此损失函数组合确保模型不仅能有效重构高维CSI幅度,还能学习和保留有意义的潜在特征。

鉴于CSI指纹间固有的差异性,以及CNN模型在定位任务中的出色表现,本文将CNN模型作为编码器的主要结构。具体而言,编码器由三个卷积层 $\text{Conv}(n_k; n_f)$ 、一个平展层和一个全连接层组成,其中 n_k 和 n_f 分别表示卷积核的数量和特征尺寸。相应地,解码器D则配置三个反卷积层 $\text{T-Conv}(n_k; n_f)$ 和一个全连接层。为了加快模型训练速度并增强其稳定性,本文在每一个Conv和T-Conv上均集成批归一化(Batch Normalization, BN)^[36]处理和LeakyReLU^[37]激活函数。

通过上述步骤获得训练好的生成器(编码器)后,将EP处的CSI指纹 \bar{A}^* 作为条件变量和随机变量一起输入到生成器中,即可获得EP处的合成CSI幅度测量。值得注意的是,当训练数据不足时会导致数据分布相对集中,使得CVAE生成的样本缺乏多样性。然而,CSI在固定位置上的稳定性较好,

即CSI数据分布同样相对集中。因此,本文重点考虑使用CSI指纹作为条件变量,提升在指纹对应位置上生成的CSI测量准确性。在获得EP处的合成CSI幅度测量后,这些数据将与真实的CSI幅度测量一同输入到特征提取模块中。

3.3 特征提取

通常,智能手机在同一位置采集的CSI幅度测量会表现出一定的稳定性,而在两个不同位置采集的CSI幅度则会表现出明显差异^[38]。这些差异极其关键,因为它们反映了不同子载波上的幅度与其相对应的位置空间的相关性。为此,本文精心设计了特征提取模块,旨在高效地提取有价值的位置特征,此模块的整体结构如图3所示。

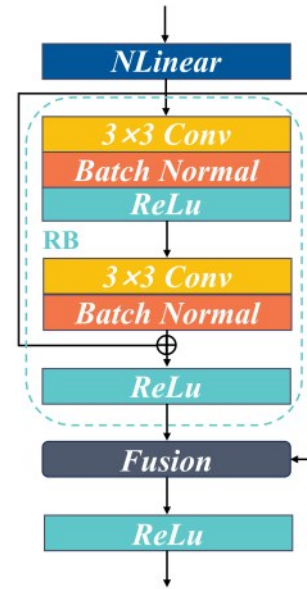


图3 特征提取模块结构

NLinear。由于在一段时间内采集的智能手机CSI幅度测量相对稳定。因此,对于稳定的CSI数据,其长程依赖可以通过更为简单的线性层进行建模,而不需要复杂的自注意力机制。基于此,采用NLinear代替结构复杂的Transformer,以提取CSI幅度中与位置相关的全局特征。与Transformer相比,NLinear具有更低的计算复杂度和更少的参数量,使其在处理高维时序数据时表现的更加高效;它通过其线性层能够直接捕捉全局趋势,并且计算复杂度为 $O(n)$,明显低于Transformer的复杂度 $O(n^2)$ ^[39]。本质上,NLinear是一个直接多步预测模型,能够高效地处理时序数据。与递归多步预测不同,它不依赖于过去的预测结果生成后续预测,从

而避免了误差累积问题。通过时间线性层对 CSI 幅度测量进行加权求和, NLinear 能够快速提取关键信息而无需复杂的非线性操作。理论上, NLinear 可以表示为:

$$X_n = W_n D_c, \quad (9)$$

其中 $W_n \in \mathbb{R}^{K' \times N_c}$, D_c 和 $X_n \in \mathbb{R}^{N_b \times K'}$ 分别为输入的 CSI 幅度测量和输出特征, N_c 代表输出特征的大小, N_b 代表处理的 CSI 样本数量。此外, 尽管 CSI 幅度分布可能随时间的推移而略微改变, 但采用 NLinear 依然能有效减少序列中的短期波动对特征提取的影响, 使模型更专注于全局趋势。具体而言, NLinear 首先减去序列中的最后一个值, 通过线性层处理剩余的序列数据, 然后再将减掉的值加回, 从而保证特征提取的稳定性。

RB. 尽管 NLinear 能够提取 CSI 数据中与位置相关的全局特征。然而, 仅依赖全局特征, 模型可能不足以区分相邻位置之间的细微信号变化差异。为此, 采用 RB^[40] 提取与 CSI 幅度相对应的位置上的局部空间特征。理论上, RB 可以被定义为:

$$X_r = \sigma(\beta W_{c2}(\sigma W_{c1}(X_n)) + X_n), \quad (10)$$

其中 X_r 为输出向量, W_{c1} 和 W_{c2} 表示两个卷积层, σ 和 β 分别代表 ReLU 激活函数和批归一化处理。与传统的 CNN 相比, RB 不仅加速模型的收敛, 而且在一定程度上缓解梯度消失问题。值得注意的是, 为了使 RB 中的卷积层能够直接处理全局特征 X_n , 本文首先将 X_n 的尺寸调整为 $N_b \times 1 \times K'$ 。然后, 将 X_r 与 X_n 的特征尺寸统一调整为 $N_b \times K'$ 。

特征融合。 本文的特征融合机制旨在将全局特征 X_n 与局部空间特征 X_r 融合成一个含有丰富判别信息的单一特征向量。传统的特征融合策略主要分为两大类: 串行融合与并行融合^[41]。具体而言, 串行融合策略通过直接拼接两个特征向量来实现, 而并行融合策略则是将这两个特征向量整合为一个向量。为了减少对智能手机内存和计算资源的占用, 本文采用并行融合策略来整合 X_n 和 X_r , 从而避免融合后的特征维度增加。

为了有效地整合 X_n 和 X_r , 本文设计一种特征融合机制, 其表达式为:

$$X_f = \alpha_1 X_n + \alpha_2 X_r + \beta, \quad (11)$$

其中 X_f 为融合后的特征, α_1 和 α_2 为权重系数; β 为偏置项, 其初始值为 0, 并随着模型训练动态调整取值。鉴于全局特征 X_n 和局部特征 X_r 同等重要,

因此本文设置 $\alpha_1 = \alpha_2 = 0.5$ 。换言之, 公式(11)通过为两个特征分配相同的权重系数来确保任意一个特征都不会主导学习过程。这种融合方式与残差网络中的跳跃连接机制类似, 它允许模型同时利用从 CSI 幅度测量中提取的整体趋势特征和局部空间特征, 以获得更加准确和健壮的定位性能。此外, 公式(11)中的偏置项与卷积层中的偏置项在作用上具有相似性, 主要目的都是调整模型输出, 以更好地匹配目标数据的分布^[42]。通过加入偏置项 β , 使模型可以更灵活地学习 CSI 幅度测量中隐藏的位置特征, 从而提升定位精度。

3.4 位置估计

在获得融合特征 X_f 后, 本文采用擅长表示输入输出之间逻辑关系的 DD 网络来估计最终结果。与传统的全连接层相比, DD 网络能够降低过拟合的风险, 增强对噪声测量的鲁棒性, 从而获得更准确和可靠的估计位置。

DD 网络^[43] 由类似神经树突的模块组成, 集成了 DD 模块和线性模块。DD 模块通过执行矩阵乘法和哈德玛乘积来模拟树突的逻辑操作, 而线性模块则负责进行线性变换。DD 模块的定义如下:

$$D^d = W^{d,d-1} D^{d-1} \odot X_f, \quad (12)$$

其中, X_f 表示 DD 网络的输入数据, D^{d-1} 和 D^d 分别代表 DD 模块的输入数据和输出数据, $W^{d,d-1}$ 表示 $d-1$ 模块到 d 模块的权重矩阵, \odot 表示哈德玛乘积。本文使用具有两个输入和一个输出的 DD 模块, 其数学表达式为:

$$\tilde{L} = W^{2,1}((W^{1,0} X_f) \odot X_f). \quad (13)$$

值得注意的是, DD 模块通过使用矩阵乘法和哈达马积(元素对元素的乘积)来提升计算效率, 它们相比于传统的非线性函数更为高效。

4 性能评估

4.1 实验设置

场景。 本文的实验在两种典型的室内场景下进行, 每种场景均包含桌椅和办公设备, 如图 4 所示。第一种场景是一个面积为 $6m \times 6m$ 的会议室, 包括 25 个 RP、5 个测试点(Test Point, TP)和 8 个 EP。第二种环境是一个总面积为 $60m^2$ 的实验室, 包括 36 个 RP、6 个 TP 和 12 个 EP。在每种场景下的每个 RP 处收集 400 个样本用于模型训练, 同样在每个 EP 处生成 400 个样本用于构造合成数据集以辅助模

型训练，并在每个TP处收集400个样本用于测试。

设备。实验中将Nexus 5和Nexus 6P智能手机作为定位目标(接收端)；华硕RT-AC86U WiFi路由器作为发射端，为实验提供必要的信号覆盖和数据传输。设置WiFi路由器的中心频率为5.745 GHz，带宽为80 MHz，对应的子载波数量为256。然后，在配备Intel Core i7-11700KF CPU (3.6 GHz)和

NVIDIA GeForce RTX3060 GPU (12 GB)的工作站上安装MATLAB R2020b和Pytorch 2.1软件处理采集到的CSI测量。

对比模型。为了验证LFLM-SCSI的定位性能，本文将其与五个主流的DL定位模型进行比较，包括基于LSTM，CNN和Transformer的变体模型。



图4 两种典型的室内真实场景及其空间布局示意图

DCNN^[16]: 该模型由四个顺序排列的卷积层构成，每个卷积层后面紧跟着一个全局池化层，最后通过一个全连接层输出定位结果；

CNN-LSTM^[26]: 该模型由三个连续的卷积层、一个LSTM层和一个全连接层组成；

FTLoc^[27]: 该模型由基于GRU的特征提取模块和自适应注意力机制组成；

SBLSTM^[28]: 该模型包括三个BiLSTM层，以及一个包含五个残差块的残差网络；

SwinLoc^[29]: 该模型以Swin Transformer为基础构建多层特征提取模块实现位置估计。

此外，为了评估LFLM-SCSI中数据增强模块的有效性，本文还将其与GPR方法进行对比。

评价指标。本文采用均方根误差(Root Mean Square Error, RMSE)、平均误差(Average Error, AE)和标准差误差(Standard Deviation Error, SDE)作为评价指标。其中，定位误差为估计位置 \hat{L} 和真实位置 L 之间的欧氏距离。

4.2 整体性能

定位精度。表1提供了LFLM-SCSI模型及其对比模型在不同场景和设备下的定位误差。实验结果表明，提出的定位模型在会议室场景和实验室场景下的平均RMSE分别为0.97m和1.26m，具有最好的定位精度。观察图5发现，LFLM-SCSI的定位误差高度集中且异常值极少，主要源于数据增强模块对多样化特征的学习，这一点在表4的模型消融实

验结果中得到了更充分的验证。SwinLoc 基于 Swin Transformer 的出色特征提取和表征能力，在定位精度上显著优于基于除 LFLM-SCSI 的传统定位模型。这与其在自然语言处理和计算机视觉等领域的出色性能表现相符。然而，LFLM-SCSI 的定位精度却优于结构复杂的 SwinLoc，在两种室内场景下的平均 RMSE 降低了 12.5%，原因如下。首先，通过分析表 4 的消融实验结果发现，特征提取模块与位置估计模块的结合仅能使 LFLM-SCSI 取得与 SwinLoc 相似的定位精度。通过引入数据增强模块，LFLM-SCSI 在训练阶段能够学习到更加多样化的位置特征，使得其箱型图展示出的定位误差范围明显小于

SwinLoc。其次，SwinLoc 模型可能需要大量的训练数据才能充分发挥出其捕捉深层特征的潜力。然而，本文提供的 CSI 样本相对有限，从而导致其定位精度未能达到预期效果。FTLoc、SBLSTM 和 DCNN 等传统模型表现较差，平均 RMSE 显著偏高，原因在于其模型的特征提取能力有限造成定位误差大且不稳定。从场景维度看，实验室环境中各模型的平均 RMSE 普遍高于会议室。例如，CNN-LSTM 在实验室（Nexus 5）环境中的 RMSE 为 1.95m，而在会议室中为 1.75m。原因在于，实验室中的人员流动和复杂设备摆放增加了定位难度，而会议室的静态环境更利于捕获特征。

表 1 不同定位模型在不同场景和设备下的定位误差(米)

定位模型	Nexus 5 会议室			Nexus 5 实验室			Nexus 6P 会议室			Nexus 6P 实验室		
	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE
CNN-LSTM	1.75	1.61	0.70	1.95	1.82	0.71	1.66	1.47	0.78	1.91	1.69	0.88
DCNN	1.64	1.48	0.72	1.82	1.77	0.42	1.50	1.49	0.16	1.91	1.89	0.26
SBLSTM	1.63	1.52	0.62	1.89	1.86	0.31	1.51	1.49	0.19	1.90	1.88	0.29
FTLoc	1.50	1.35	0.66	1.82	1.65	0.78	1.53	1.51	0.18	1.87	1.67	0.84
SwinLoc	1.28	1.03	0.77	1.47	1.38	0.52	1.07	1.01	0.38	1.33	1.23	0.50
LFLM-SCSI	1.10	0.98	0.51	1.25	1.21	0.34	0.83	0.73	0.40	1.27	1.19	0.45

物理资源。所有定位模型的物理资源消耗如表 2 所示，主要包括浮点计算、内存占用和模型参数数量。这些指标对于评估模型的计算复杂度和硬件资源需求具有重要意义。表中的结果基于 torchinfo^① 工具输出。实验结果表明，SwinLoc 以 27.21 M 的参数规模居首，LFLM-SCSI 通过结构优化将参数控制在 21.73 M，虽显著低 SwinLoc 但仍高于其他模型；DCNN 则以 0.01 M 的参数数量最少，SBLSTM (1.19 M)、FTLoc (0.20 M) 和 CNN-LSTM (12.36 M) 参数规模依次递增。内存占用趋势与参数数量高度一致，SwinLoc 因结构复杂与高参数需求导致内存占用过高 (534.74 MB)，LFLM-SCSI 通过优化将内存需求降至 280.46 MB，而 DCNN 仅需 2.44 MB 内存，凸显其硬件资源友好性；SBLSTM (154.72 MB)、CNN-LSTM (110.85 MB) 和 FTLoc (24.98 MB) 的内存占用则随模型复杂度阶梯式上升。浮点计算量方面，CNN-LSTM (8.88 G) 与 LFLM-SCSI (7.76 G) 因融合多模块特征计算需求较高，SwinLoc (1.09 G) 虽参数与内存占优，但其 Transformer 结构的注意力机制计

算特性导致浮点运算量低于预期；DCNN (12.51 M) 凭借极简结构实现最低浮点计算消耗，与 FTLoc (588 M) 构成中低计算复杂度梯队。

时间成本。表 3 提供了不同定位模型在不同场景和设备下的时间成本，具体包括训练时间和推理时间。其中，训练时间指模型训练过程中单个 epoch 的执行时间，而推理时间则是指完成单次定位任务所需时间。这两个指标各自反映模型在离线训练阶段对时间资源的消耗和定位阶段可能出现的定位时延。实验结果表明，物理资源与时间成本存在一定的关联。DCNN 因结构极简、参数少且无复杂模块，训练时长最短，在 23 秒~45 秒之间，至少比其它定位模型明显缩短了 6~7 倍；FTLoc 结构复杂度中等，特征融合策略简单，训练耗时在 53 秒~73 秒之间；SBLSTM 基于 LSTM 时序建模，需处理序列依赖，训练耗时中等；CNN-LSTM 融合空间与时序特征，计算量增加，训练耗时与 SBLSTM 数量级相同；SwinLoc 结构复杂 (参数 27.21 M)，Swin Transformer 的分层特征提取与窗口注意力计算使计算量极大，训练时间最长，在 1200 秒~1800

①<https://github.com/TylerYep/torchinfo>

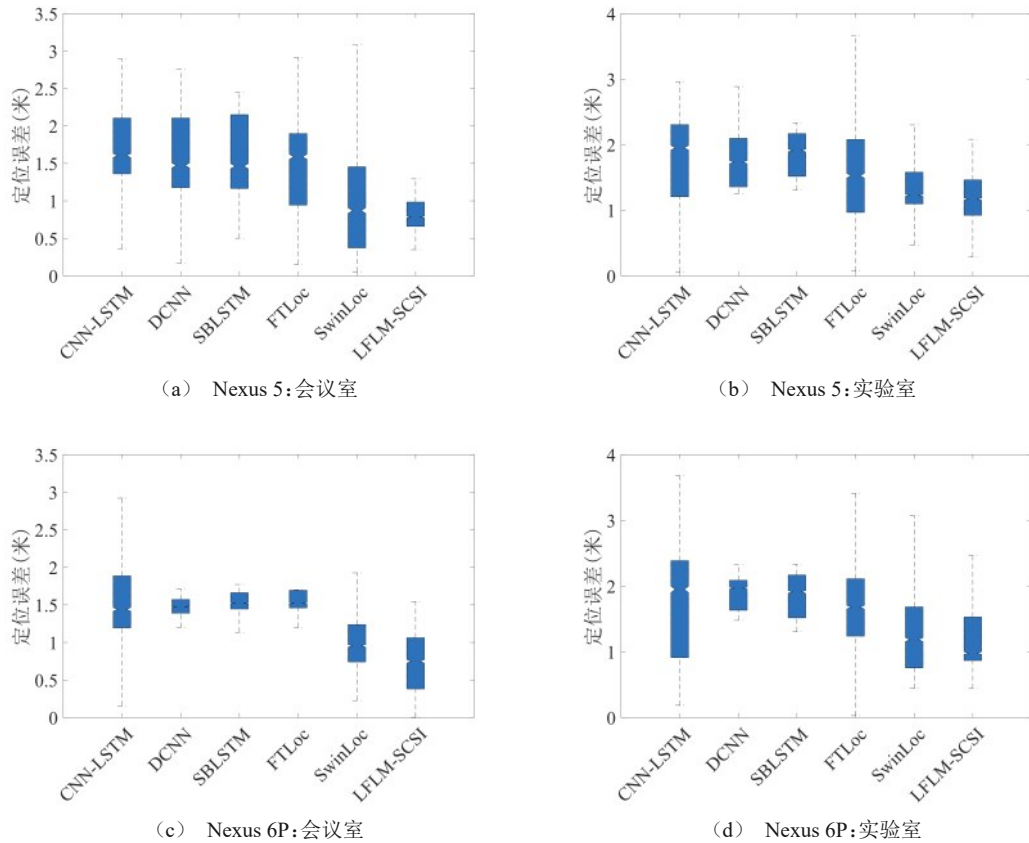


图5 不同定位模型的定位误差箱型图

秒之间；LFLM-SCSI因数据增强模块及合成数据扩增，训练耗时在420秒~550秒之间，约为SwinLoc的1/3，但结构优化使其参数低于SwinLoc。

综上所述，高参数、高内存占用的定位模型（如SwinLoc）训练耗时显著增加，而DCNN通过极低的资源消耗实现了最优时间效率，但定位精度受限于特征提取能力不足。LFLM-SCSI则在资源消耗与定位精度间取得平衡，其优化后的参数规模与内存占用能够避免如SwinLoc的资源冗余问题，为高精度定位场景的硬件部署提供了可行性。

表2 不同定位模型的资源消耗

模型	浮点计算	内存占用	参数数量
CNN-LSTM	8.88G	110.85MB	12.36M
DCNN	12.51M	2.44MB	0.01M
SBLSTM	3.54G	154.72MB	1.19M
FTLoc	588M	24.98M	0.20M
SwinLoc	1.09G	534.74MB	27.21M
LFLM-SCSI	7.76G	280.46MB	21.73M

4.3 消融实验

定位精度。本文通过消融实验来评估LFLM-SCSI模型中关键组件的有效性，这些关键组件包

表3 不同定位模型的训练时间和推理时间(秒)

模型	Nexus 5 会议室		Nexus 5 实验室		Nexus 6P 会议室		Nexus 6P 实验室	
	训练时间	推理时间	训练时间	推理时间	训练时间	推理时间	训练时间	推理时间
CNN-LSTM	280.83	0.51	292.26	0.58	217.52	0.52	298.42	0.59
DCNN	23.93	0.05	42.78	0.06	27.16	0.06	32.21	0.06
SBLSTM	168.93	0.44	252.18	0.55	171.39	0.45	249.60	0.54
FTLoc	53.09	0.11	72.69	0.12	64.69	0.11	73.71	0.12
SwinLoc	1261.71	2.83	1778.96	3.23	1202.25	2.59	1708.98	3.09
LFLM-SCSI	417.11	0.07	530.07	0.08	419.37	0.06	545.52	0.07

括:NLinear(N)、RB(R)、特征融合(F)、DD网络(D)、PCA降维(P)和数据增强(A)。表4和图6展示了LFLM-SCSI中不同组件的组合对定位精度的影响。实验结果表明,提出的定位模型在不同设备和场景下的定位精度,随着更多组件的集成而逐步提升。从基础组件N开始,逐步添加R、D、F、A组件,使得定位误差逐渐降低。在集成所有组件

(N+R+D+F+A)后,会议室场景和实验室场景下的平均RMSE分别减少了36.6%和25.4%,表现出了较好的定位精度。然而,在加入P组件后,模型在两种场景下的平均RMSE反而出现增长。原因在于PCA降维可能导致CSI幅度测量中的有效位置特征被忽略。

表4 LFLM-SCSI中不同组件的定位误差(米)

组件	Nexus 5会议室			Nexus 5实验室			Nexus 6P会议室			Nexus 6P实验室		
	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE
N	1.60	1.50	0.53	1.72	1.57	0.69	1.45	1.23	0.77	1.65	1.48	0.73
N+R	1.51	1.30	0.77	1.55	1.45	0.57	1.41	1.34	0.45	1.58	1.42	0.70
N+R+D	1.36	1.13	0.75	1.49	1.40	0.49	1.29	1.23	0.38	1.52	1.37	0.66
N+R+D+F	1.25	1.11	0.58	1.38	1.31	0.42	0.91	0.82	0.39	1.44	1.34	0.52
N+R+D+F+P	1.46	1.30	0.68	1.56	1.43	0.62	1.32	1.10	0.73	1.57	1.44	0.65
N+R+D+F+A	1.10	0.98	0.51	1.25	1.21	0.34	0.83	0.73	0.40	1.27	1.19	0.45

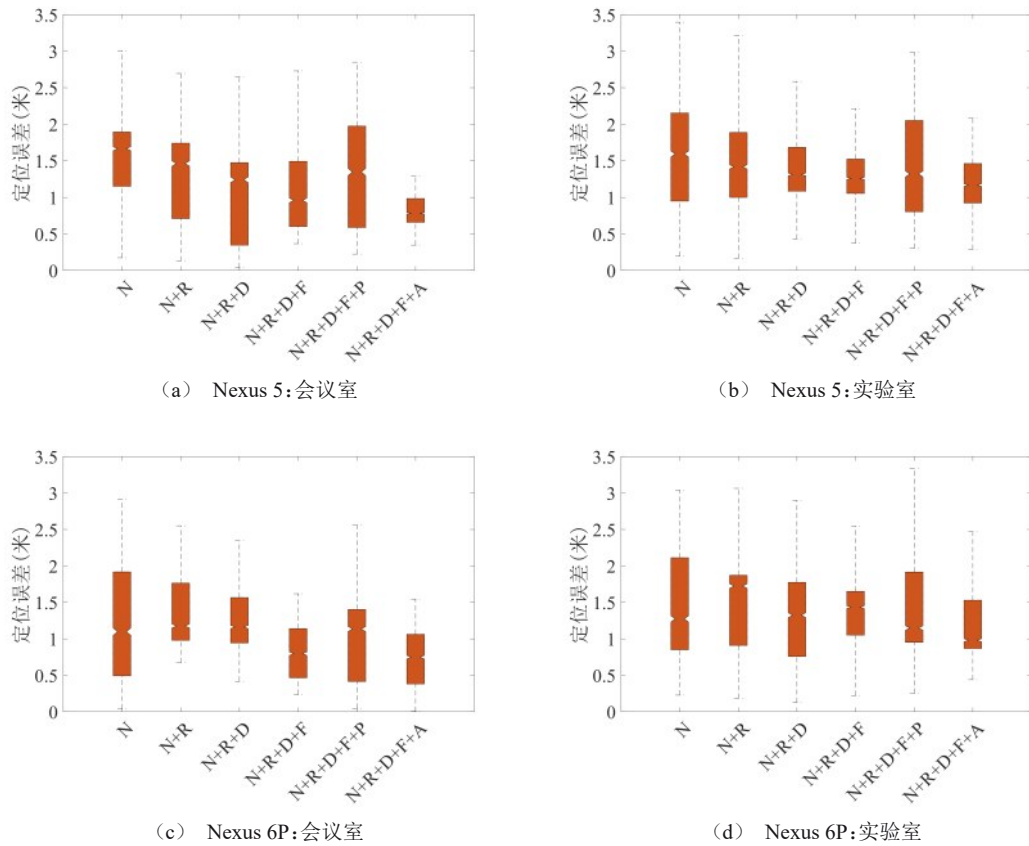


图6 LFLM-SCSI中不同组件的定位误差箱型图

为了进一步评估各组件对于模型定位性能的影响,在表5和表6详细列举了LFLM-SCSI中各组件在不同场景和设备下的时间成本和物理资源消耗情况。

时间成本。表5展示了LFLM-SCSI中各组件在不同室内环境和设备下的训练时间和推理时间。随着更多组件的加入,模型的训练时间逐渐增加,并在特征增强模块(A)加入后发生明显增长。然

而，各组件的推理时间变化不超过0.04秒。基础组件NLinear (N) 具有最短的训练和推理时间，它在会议室场景和实验室场景下的平均训练时间分别为14.75秒和21.56秒，平均推理时间分别为0.03秒和0.04秒。加入R、D、F、A组件后，会议室场景和实验室场景下的平均训练时间分别增长了418.24秒和537.79秒，但平均推理用时仅增长0.04秒。这

表明新增组件对训练时间产生明显影响，而对推理时间的影响较小。因此，本文提出的定位模型在实际应用中仍能保持较快的响应时间。值得注意的是，当加入P组件后，训练时间显著下降，在会议室场景和实验室场景下的平均训练时间分别降至18.03秒和26.28秒，推理时间也保持在较低水平，平均推理用时分别为0.04秒和0.05秒。

表5 LFLM-SCSI中不同组件的训练时间和推理时间(秒)

组件	Nexus 5会议室		Nexus 5实验室		Nexus 6P会议室		Nexus 6P实验室	
	训练时间	推理时间	训练时间	推理时间	训练时间	推理时间	训练时间	推理时间
N	14.99	0.03	22.38	0.04	14.51	0.03	20.79	0.04
N+R	24.05	0.05	33.75	0.06	21.84	0.05	31.23	0.06
N+R+D	29.42	0.07	38.62	0.07	25.13	0.06	37.26	0.07
N+R+D+F	33.21	0.07	39.99	0.08	26.13	0.06	42.08	0.07
N+R+D+F+P	17.93	0.04	26.96	0.04	18.12	0.04	25.59	0.05
N+R+D+F+A	417.11	0.07	530.07	0.08	419.37	0.06	545.52	0.07

表6 LFLM-SCSI不同组件的资源消耗

组件	浮点计算	内存占用	参数数量
N	0.86M	0.17MB	0.03M
N+R	0.89M	0.30MB	0.03M
N+R+D	1.41M	0.40MB	0.04M
N+R+D+F	1.41M	0.40MB	0.04M
N+R+D+F+P	0.77M	0.30MB	0.03M
N+R+D+F+A	7.76G	280.46MB	21.73M

物理资源。表6总结了LFLM-SCSI模型中各组件的资源占用情况。实验结果表明，基础组件N具有最低的资源消耗，其浮点计算量、内存使用和参数数量分别为0.86M、0.17MB和0.03M，非常适合在资源受限的环境中使用。加入组件R后，资源消耗略有增加，浮点计算量和内存使用分别增加了0.03M和0.13MB。进一步添加组件D和F后，模型的浮点计算量、内存使用和参数数量也逐步增长至1.41M、0.40MB和0.04M。值得注意的是，在加入组件F后，模型的资源占用较为稳定，没有出现增长。产生这一现象的原因在于，组件F的加入并未引入额外的参数，因此不会增加额外的内存占用。而在加入组件P后，资源消耗呈下降趋势，浮点计算从1.41M降至0.77M，内存占用从0.40MB回落至0.30MB，参数数量也恢复至0.03M，这表明组件P的引入有效降低了资源消耗，但定位精度也随之下降。此外，加入组件A后，模型的资源占用增长十分明显。然而，组件A仅在模型训练过程中用

于提供合成数据集，而不直接参与位置推理过程。因此，模型在实际推理过程中的资源占用应当与组合(N+R+D+F)的占用相近。

CVAE有效性。为了验证通过数据增强模块生成的智能手机CSI幅度测量的有效性，使用特征提取和位置估计模块，对CVAE和GPR构造的合成数据集进行评估。图7和表7展示了在不同情况下，使用GPR和GPR-CVAE提供的合成数据集后所取得的定位误差。实验结果表明，无论处于何种情况，GPR-CVAE的平均RMSE、AE和SDE均优于GPR。具体而言，在会议室场景中，GPR-CVAE的平均RMSE、AE和SDE为0.97m、0.86m和0.46m，与GPR相比各项指标依次减少了0.07m、0.07m和0.02m。GPR-CVAE和GPR在实验室场景下的定位性能对比与会议室场景中类似。虽然GPR-CVAE在实验室场景中的AE和SDE指标会略有增长，但其整体定位精度提升明显。GPR-CVAE模型之所以能取得比单独使用GPR更好的效果，主要得益于CVAE能生成更多样化和具有代表性的CSI幅度测量。这些数据能够更加全面地反映环境中信号随距离变化的特征。将这些合成数据引入定位模型的训练过程中，可以有效扩大模型对位置空间的学习范围，从而显著提升定位精度和鲁棒性。

核函数对比。由于可靠的CSI指纹是在EP处准确生成CSI幅度测量的基础，因此选择适当的核函数确保CSI指纹估计精度显得尤为重要。为此，

本文评估了包括RBF、Matern、RQ及其组合核在估计CSI指纹方面的效果。具体而言，从实验室中和会议室中分别随机选取6个RP和5个RP的CSI指纹用于测试，而将其他RP的CSI指纹用于训练基于不同核函数的GPR模型。表8展示了在不同场景下，应用不同核函数组合的GPR模型取得的CSI指纹误差估计结果。由于对CSI测量进行了标准化处理，因此将表8中的误差结果保留至小数点后四位，以更加直观地展示不同核函数组合对CSI指纹估计准确性的影响。实验结果表明，在会议室和实

验室两种场景中，采用不同核函数组合的CSI指纹估计误差的确存在明显差异。值得注意的是，Matern+RQ核函数组合在两种场景下的表现最佳，其RMSE、AE和SDE均为最小值，表明该组合在CSI指纹估计中具有最好的准确性。这主要得益于Matern核和RQ核的互补能力，这两种核函数的组合有助于准确捕捉室内空间中CSI幅度测量与距离之间的映射关系。因此，选择基于Matern+RQ核函数的GPR模型估计EP处的CSI指纹信息。

表7 采用GPR和CVAE合成数据的定位误差(米)

模型	Nexus 5会议室			Nexus 5实验室			Nexus 6P会议室			Nexus 6P实验室		
	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE
GPR	1.21	1.08	0.54	1.34	1.19	0.63	0.87	0.77	0.41	1.35	1.31	0.36
CVAE	1.10	0.98	0.51	1.25	1.21	0.34	0.83	0.73	0.40	1.27	1.19	0.45

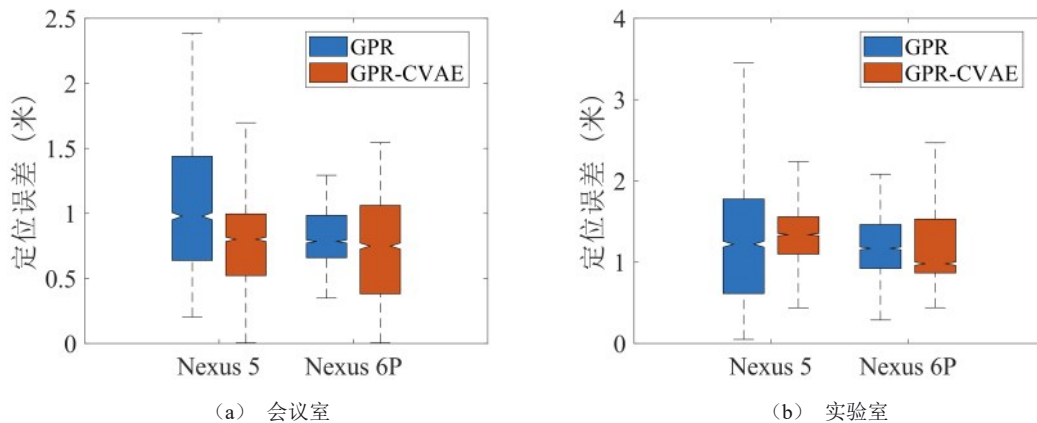


图7 采用GPR和CVAE合成数据的定位误差箱型图

表8 不同核函数下的CSI指纹估计误差

核函数	Nexus 5会议室			Nexus 5实验室			Nexus 6P会议室			Nexus 6P实验室		
	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE	RMSE	AE	SDE
RBF	0.3035	0.2501	0.1720	0.2995	0.2424	0.1759	0.3039	0.2438	0.1813	0.3086	0.2506	0.1801
RQ	0.3067	0.2504	0.1771	0.3292	0.2695	0.1891	0.3243	0.2627	0.1902	0.3119	0.2538	0.1813
Matern	0.3061	0.2497	0.1770	0.2976	0.2428	0.1714	0.3037	0.2481	0.1750	0.2970	0.2399	0.1751
RBF+RQ	0.2739	0.2206	0.1622	0.2972	0.2425	0.1718	0.3038	0.2438	0.1813	0.3088	0.2508	0.1802
Matern+RQ	0.2705	0.2179	0.1602	0.2837	0.2280	0.1688	0.3037	0.2473	0.1744	0.2959	0.2387	0.1748
Matern+RBF	0.2897	0.2371	0.1665	0.2978	0.2429	0.1727	0.3029	0.2474	0.1747	0.2979	0.2409	0.1753

5 结束语

本文提出了一种基于智能手机CSI的轻量级指纹定位模型LFLM-SCSI，该模型由数据增强模块、特征提取模块和位置估计模块组成。数据增强模块首先利用GPR计算EP处的CSI指纹作为条件变量，

然后将其和随机变量一起输入到训练好的解码器中，生成可靠且多样化的CSI幅度测量。获得合成CSI数据后，特征提取模块将合成数据与真实数据相结合，利用线性层NLinear和残差块RB提取CSI幅度测量中与空间位置相关的整体趋势特征和局部空间特征。同时，利用特征融合机制在不增加模型

的参数数量的情况下，将上述特征有效融合。最后，采用基于DD网络的位置估计模块估计位置。本文在两个真实的室内场景下对LFLM-SCSI的性能进行验证。实验结果表明，与表现最好基线模型SwinLoc相比，LFLM-SCSI不仅展现较高的定位精度，还具有较低的资源占用和较短的定位用时。

参考文献：

- [1] Farahsari P S, Farahzadi A, Rezazadeh J, et al. A survey on indoor positioning systems for iot-based applications. *IEEE Internet of Things Journal*, 2022, 9(10):7680 - 7699.
- [2] Chen X, Li H, Zhou C, et al. Fidora: Robust wifi-based indoor localization via unsupervised domain adaptation. *IEEE Internet of Things Journal*, 2022, 9(12):9872 - 9888.
- [3] Zheng S, Zhou Z, Zhang Q, et al. A factor graph based indoor localization approach for healthcare//Proceedings of the 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Istanbul and Turkey, 2023:3257 - 3264.
- [4] Zhou Y, Huang H, Yuan S, et al. Metafi+ : wifi-enabled transformer-based human pose estimation for metaverse avatar simulation. *IEEE Internet of Things Journal*, 2023, 10(16):14128-14136.
- [5] Zhu X, Qu W, Qiu T, et al. Indoor intelligent fingerprint-based localization: Principles, approaches and challenges. *IEEE Communications Surveys & Tutorials*, 2020, 22(4):2634 - 2657.
- [6] He S and Chan S H G. Wifi fingerprint-based indoor positioning: recent advances and comparisons. *IEEE Communications Surveys & Tutorials*, 2016, 18(1):466-490.
- [7] Wu K, Xiao J, Yi Y, et al. FILA: Fine-grained indoor localization [C]//2012 Proceedings IEEE INFOCOM. IEEE, 2012: 2210-2218.
- [8] Kotaru M, Joshi K, Bharadia D, et al. Spotfi: Decimeter level localization using wifi//Proceedings of the 2015 ACM conference on special interest group on data communication. London, UK, 2015: 269-282.
- [9] Soltanaghaei E, Kalyanaraman A, Whitehouse K. Multipath triangulation: Decimeter-level wifi localization and orientation with a single unaided receiver//Proceedings of the 16th annual international conference on mobile systems, applications, and services. Munich, Germany 2018: 376-388.
- [10] Xie Y, Xiong J, Li M, et al. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking[C]//The 25th Annual International Conference on Mobile Computing and Networking. 2019: 1-16.
- [11] Li Y, Yan K. Indoor localization based on radio and sensor measurements. *IEEE Sensors Journal*, 2021, 21(22): 25090-25097.
- [12] Sobehy A, Renault É and Mühlethaler P. Csi-mimo: k-nearest neighbor applied to indoor localization//Proceedings of the 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020:1-6.
- [13] Dang X, Si X, Hao Z, et al. A novel passive indoor localization method by fusion CSI amplitude and phase information[J]. *Sensors*, 2019, 19(4): 875.
- [14] Xiao J, Wu K, Yi Y, et al. Pilot: Passive device-free indoor localization using channel state information[C]//2013 IEEE 33rd International Conference on Distributed Computing Systems. IEEE, 2013: 236-245.
- [15] Song X, Fan X, He X, et al. Cnnloc: deep-learning based indoor localization with wifi fingerprinting//Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation, Paris, 2019:589 - 595.
- [16] Wang X, Wang X, and Mao S. Deep convolutional neural networks for indoor localization with csi images. *IEEE Transactions on Network Science and Engineering*, 2020, 7(1):316 - 327.
- [17] Choi J S, Lee W H, Lee J H, et al. Deep learning based NLOS identification with commodity WLAN devices. *IEEE Transactions on Vehicular Technology*, 2017, 67(4): 3295-3303.
- [18] Chen C, Chen Y, Han Y, et al. Achieving centimeter-accuracy indoor localization on WiFi platforms: A frequency hopping approach[J]. *IEEE Internet of Things Journal*, 2016, 4(1): 111-121.
- [19] Song Q, Guo S, Liu X, et al. Csi amplitude fingerprinting-based nb-iot indoor localization. *IEEE Internet of Things Journal*, 2018, 5(3):1494 - 1504.
- [20] Sanam T F and Godrich H. Fuseloc: A cca based information fusion for indoor localization using csi phase and amplitude of wifi signals//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. London, UK, 2019:7565 - 7569.
- [21] Zhou R, Lu X, Zhao P, et al. Device-free presence detection and localization with svm and csi fingerprinting. *IEEE Sensors Journal*, 2017, 17(23):7990 - 7999.
- [22] Wang Y, Xiu C, Zhang X, et al. Wifi indoor localization with csi fingerprinting-based random forest. *Sensors*, 2018, 18(9):1-23.
- [23] Wang X, Gao L, Mao S, et al. Csi-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 2017, 66(1):763 - 776.
- [24] Wang X, Gao L, Mao S. CSI phase fingerprinting for indoor localization with a deep learning approach[J]. *IEEE Internet of Things Journal*, 2016, 3(6): 1113-1123.
- [25] Wang X, Gao L, Mao S. BiLoc: Bi-modal deep learning for indoor localization with commodity 5GHz WiFi[J]. *IEEE access*, 2017, 5: 4209-4220.
- [26] Hoang M T, Yuen B, Ren K, et al. A cnn-lstm quantifier for single access point csi indoor localization. *ArXiv e-prints*, 2020:1 - 10.
- [27] M. Liu, X. Liao, Z. Gao, et al. FT-Loc: A Fine-Grained Temporal Features-Based Fusion Network for Indoor Localization[J]. *IEEE Internet of Things Journal*, 2024, 11(3):4324-4334.

- [28] Fan S, Wu Y, Han C, et al. A structured bidirectional lstm deep learning method for 3d terahertz indoor localization//Proceedings of the 2020 IEEE Conference on Computer Communications. Virtual Conference, 2020:2381 - 2390.
- [29] Xu X, Zhu F, Han S, et al. Swin-loc: transformer-based csi fingerprinting indoor localization with mimo isac system[J]. IEEE Transactions on Vehicular Technology, 2024:1-16.
- [30] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Virtual Conference, 2021:10012-10022.
- [31] Wang J, Li F, Ma X, et al. Lightweight device-free wireless sensing using information-for-complexity strategy[J]. IEEE Transactions on Industrial Informatics, 2022, 19(4): 6117-6126.
- [32] Yousefi S, Narui H, Dayal S, et al. A survey on behavior recognition using WiFi channel state information. IEEE Communications Magazine, 2017, 55(10): 98-104.
- [33] Sohn K, Lee H, Yan X. Learning structured output representation using deep conditional generative models//Proceedings of Advances in neural information processing systems. Montreal, Canada, 2015:1-9.
- [34] Rasmussen C E. Gaussian processes in machine learning//Advanced Lectures on Machine Learning. Berlin, Germany, 2003: 63-71.
- [35] Li X, Li T, Ma J. Gaussian process mixture model for flexible prediction of coal mine gas concentration under noisy input prediction strategy. Signal Processing, 2021, 37(11):2031 - 2040.
- 李晓燕, 李弢, 马尽文. 高斯过程混合模型在含噪输入预测策略下的煤矿瓦斯浓度柔性预测. 信号处理, 2021, 37(11): 2031-2040.
- [36] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift//International conference on machine learning. Lille, France, 2015: 448-456.
- [37] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models//International conference on machine learning. Atlanta, USA, 2013:1-6.
- [38] Gao Z, Gao Y, Wang S, et al. CRISLoc: Reconstructable CSI fingerprinting for indoor smartphone localization. IEEE Internet of Things Journal, 2020, 8(5): 3422-3437.
- [39] Zeng A, Chen M, Zhang L, et al. Are transformers effective for time series forecasting//Proceedings of the AAAI conference on artificial intelligence. Washington, USA, 2023, 37(9): 11121-11128.
- [40] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, USA, 2016: 770-778.
- [41] Yang J, Yang J, Zhang D, et al. Feature fusion: parallel strategy vs. serial strategy. Pattern recognition, 2003, 36(6): 1369-1381.
- [42] Han K, Wang Y, Tian Q, et al. Ghostnet: More features from cheap operations[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1580-1589.
- [43] Liu G, Wang J. Dendrite net: A white-box module for classification, regression, and system identification. IEEE Transactions on Cybernetics, 2021, 52(12): 13774-13787.