



改进MixNet的CSI图像指纹室内定位方法

龙良, 王小鹏, 李岗, 王江

Improved MixNet for indoor localization using CSI image fingerprints

LONG Liang, WANG Xiaopeng, LI Gang, and WANG Jiang

引用本文:

龙良, 王小鹏, 李岗, 等. 改进MixNet的CSI图像指纹室内定位方法[J]. *全球定位系统*, 2024, 49(3): 57–64. DOI: [10.12265/j.gnss.2023198](https://doi.org/10.12265/j.gnss.2023198)

LONG Liang, WANG Xiaopeng, LI Gang, et al. Improved MixNet for indoor localization using CSI image fingerprints[J]. *Gnss World of China*, 2024, 49(3): 57–64. DOI: [10.12265/j.gnss.2023198](https://doi.org/10.12265/j.gnss.2023198)

在线阅读 View online: <https://doi.org/10.12265/j.gnss.2023198>

您可能感兴趣的其他文章

Articles you may be interested in

[基于PCA-SMO的CSI指纹定位方法](#)

Fingerprint positioning method of CSI based on PCA-SMO

全球定位系统. 2021, 46(1): 13–19

[基于人工鱼群优化的BP神经网络WiFi指纹室内定位方法](#)

WiFi fingerprint indoor location method with BP neural network based on improved artificial fish swarm optimization algorithm

全球定位系统. 2020, 45(1): 82–87

[一种基于智能手机四向RSS指纹的室内定位方法](#)

An indoor location method based on smart phone's RSS fingerprint in four directions

全球定位系统. 2021, 46(5): 48–54

[基于MIMU / WiFi指纹组合室内定位系统的研究](#)

Research on Indoor Location System Based on MIMU / WiFi Fingerprint Combination

全球定位系统. 2018, 43(5): 105–111

[融合惯导信息的单目视觉室内定位方法](#)

Indoor location method of monocular vision fusing inertial navigation information

全球定位系统. 2019, 44(3): 74–80

[基于改进AP选择的融合随机森林室内定位算法](#)

Indoor location algorithm based on improved AP selection and random forest fusion

全球定位系统. 2021, 46(5): 33–38



关注微信公众号, 获得更多资讯信息

DOI: 10.12265/j.gnss.2023198

改进 MixNet 的 CSI 图像指纹室内定位方法

龙良, 王小鹏, 李岗, 王江

(兰州交通大学电子与信息工程学院, 兰州 730070)

摘要: 为了提升信道状态信息 (channel state information, CSI) 指纹室内定位的性能, 提出了一种改进 MixNet 的 CSI 图像指纹室内定位方法. 在离线阶段, 通过选择定位参考点 (reference point, RP) 处信号强度指示 (received signal strength indication, RSSI) 最强的 3 个接入点 (access point, AP), 提取其 CSI 数据并转换为图像; 然后利用改进的 MixNet 模型对图像进行训练并保存模型. 其中改进的 MixNet 引入了坐标注意力 (coordinate attention, CA) 和残差连接. 首先, 将 MixNet-s 中的 SE (squeeze-and-excitation) 注意力替换为 CA, 以增强网络的信息表示能力并更精确地获取 CSI 图像指纹特征. 其次, 根据 MixNet-s 模型的特点构建残差连接, 以增强网络的表示能力并防止过拟合. 最后, 通过减小网络深度确保所有网络层得到充分训练; 在线阶段, 采集目标设备的 CSI 数据并转换为图像, 输入已训练好的改进 MixNet 模型 (命名为 MixNet-CA); 最后利用加权质心算法根据模型输出的概率值估计目标设备的最终位置. 提出方法在室内环境中进行了验证, 达到了 0.362 0 m 的平均定位误差.

关键词: MixNet; 坐标注意力; Wi-Fi 指纹室内定位; 信道状态信息图像; 残差连接

中图分类号: P228 **文献标志码:** A **文章编号:** 1008-9268(2024)03-0057-08

0 引言

室内定位在现代生活的各个方面都发挥着重要作用, 包括室内导航、智能建筑管理和个性化服务^[1-3]. 在室内定位技术中, 基于 Wi-Fi 的方法因其成本低、易于部署和可扩展性高而备受关注.

目前, 许多 Wi-Fi 室内定位方法利用接收信号强度 (received signal strength, RSS)^[4-7] 进行定位. 然而, 在复杂的室内环境中, 这种方法容易受到诸如多径效应、阴影衰落和信号干扰等因素的影响, 从而限制了定位的准确性和可靠性. 相比之下, 信道状态信息 (channel state information, CSI) 提供了更详细的物理层信息, 例如多通道子载波的相位和幅度. 这些细节能更准确地反映信号在传输过程中的变化. 然而, 由于 CSI 数据不直接提供有关无线信道的有用信息, 研究人员提出使用 CSI 图像作为室内定位的指纹. 通过将 CSI 数据转换为图像, 研究人员可以全面利用 CSI 中包含的信息, 避免主观的特征选择和预处理步

骤. 鉴于深度学习网络在图像特征提取方面的出色能力, 研究者选择采用深度学习网络来提取 CSI 图像的特征^[8-12]. 方法 ConFi^[8] 中使用天线的振幅信息构建 CSI 特征图像, 使用卷积神经网络 (convolutional neural network, CNN) 作为定位模型, 实现了 1.37 m 的平均定位误差. 方法 ILCL^[9] 中使用来自三个天线的相位信息构建尺寸为 30×30×3 的 RGB 特征图像, 使用 CNN 作为定位模型, 实现了 1.28 m 的平均定位误差. 方法 CiFi^[10] 中利用相邻天线之间的相位差信息构建 CSI 特征图像, 生成尺寸为 60×60 的单通道图像, 使用深度卷积神经网络 (deep convolutional neural network, DCNN) 作为定位模型, 实现了 1.78 m 的平均定位误差. 文献 [11] 结合了 CSI 的振幅和相位信息, 构建了尺寸为 114×114×3 的 RGB 特征图像, 采用轻量级 CNN (ShuffleNet V2) 作为定位模型, 实现了 1.57 m 的平均定位误差. 文献 [12] 联合 CSI 的幅度差和相位差构建指纹, 采用 CNN 作为定位模型, 实现了 0.48 m 的平均定位误差. 然而, 上述文献中的深

收稿日期: 2023-10-16

资助项目: 兰州市科技计划项目 (2023-3-104); 甘肃省高校产业支撑计划项目 (2023CYZC-40); 甘肃省优秀研究生“创新之星”项目 (2023CXZY-546)

通信作者: 王小鹏 E-mail: wangxp1969@sina.com

度学习模型在定位过程中需要大量的计算资源,并且其指纹特征提取能力有限.此外,上述文献中的定位方法仅利用了来自单个接入点(access point, AP)的CSI,未全面考虑不同AP路径之间的指纹差异.特别是在复杂的室内环境中,多径效应的影响可能会扭曲来自1个AP的CSI数据,但来自其他AP的数据可以进行补偿,从而显著减轻多径效应.

为了使CSI图像指纹能够全面考虑不同AP路径之间的指纹差异,同时确保深度学习模型更轻量化且具有更强的特征提取能力,提出了一种改进MixNet的CSI图像指纹室内定位方法.首先,通过联合利用选定的多个AP的振幅和相位差信息构建CSI图像,以全面描述室内环境中的信号特征.此外,引入了改进的MixNet网络,对MixNet-s进行改进,加入坐标注意力(coordinate attention, CA)^[13]和残差连接^[14]以更精确地获取CSI图像指纹特征并防止过拟合.随后,对网络深度进行了适当调整.

1 CSI

CSI描述了信号的传播,并揭示了信号传输过程中的散射和衰减效应.它提供了有关通信链路信道特性的信息,允许对信道进行估计.在正交频分复用(orthogonal frequency division multiplexing, OFDM)系统中,经过多径信道后的接收信号可以表示为

$$Y = HX + N \quad (1)$$

式中: Y 为接收向量; X 为发送向量; H 为信道矩阵; N 为加性白噪声(white noise, WN)向量.

信道矩阵 H 记录了无线信道的特性,包括CSI

的振幅和相位信息.商用无线网卡(如Intel 5300)集成了支持IEEE 802.11n标准的OFDM技术,能够直接提取CSI.利用Intel 5300提取30个子载波的CSI,其中第 i 个子载波的CSI可以表示为

$$H_i = |H_i| \exp(j\angle H_i) \quad (2)$$

式中: $|H_i|$ 为子载波的振幅; $\angle H_i$ 为子载波的相位.

2 MixNet-CA

2.1 MixNet-s

MixNet是由Google团队的Tan和Le在2019年提出的轻量级卷积神经网络^[15].该网络采用了一种基于初始卷积和深度可分卷积的新型卷积方法,称为混合深度卷积(mixed depthwise convolution, MixConv).

在深度可分离卷积中,传统的卷积操作被分为两个过程:深度卷积和逐点卷积.深度卷积将具有 m 个通道的图像根据通道数分为 m 组,然后使用相同的卷积核对每个组进行卷积.逐点卷积使用 1×1 卷积核,将深度卷积输出的特征图中每个通道相同位置的特征信息进行加权组合.MixConv将Inception与深度卷积结合起来.MixConv根据通道数将具有 m 个通道的图像分为 n 组,并使用不同的卷积核对每个组进行卷积^[15].MixNet-s中MixConv的输出并非直接传入逐点卷积过程,而是传入SE(squeeze-and-excitation)注意机制.逐点卷积的输入是SE机制的输出.MixConv、SE和逐点卷积按顺序封装在MixNetBlock中.MixNetBlock的结构如图1所示.

MixNet的骨干网由3个 3×3 的常规卷积和13个MixNetBlock组成.MixNet-s的主要参数如表1所示.

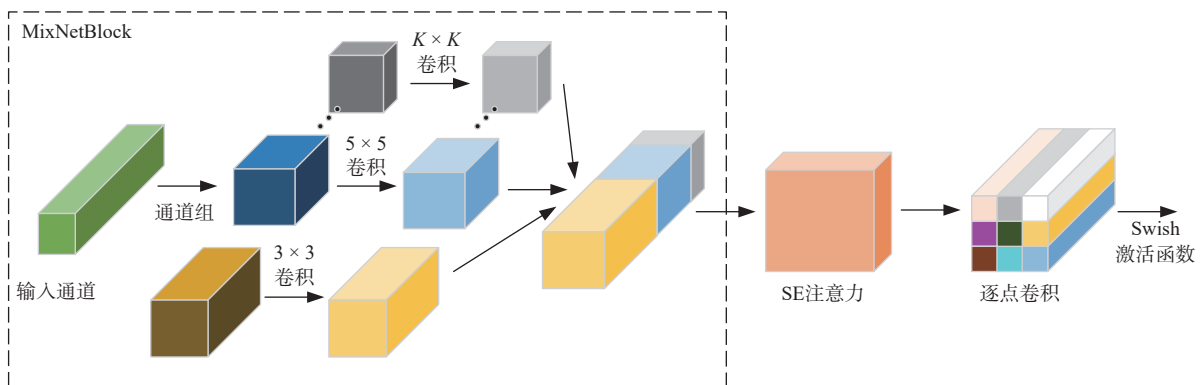


图1 MixNetBlock 结构

表1 MixNet-s 的主要参数

t	c	n	k	s	a	se
1	14	1	[3]	1	ReLU	None
6	24	1	[3]	2	ReLU	None
3	24	1	[3]	1	ReLU	None
6	40	1	[3 5 7]	2	Swish	0.5
6	40	3	[3 5]	1	Swish	0.5
6	80	1	[3 5 7]	2	Swish	0.25
6	80	2	[3 5]	1	Swish	0.25
6	120	1	[3 5 7]	1	Swish	0.5
3	120	2	[3 5 7 9]	1	Swish	0.5
6	200	1	[3 5 7 9 11]	2	Swish	0.5

注: t =扩展比率, c =输出通道, n =组数, k =MixConv 卷积核分布, s =步幅, a =激活函数, se =SE权重.

2.2 CA

提出方法将 MixNet-s 中原有的 SE 机制替换为 CA. SE 仅关注特征通道之间的信息关联, 而 CA 还关注特征空间维度上的信息关联. CA 通过学习自适应信道权值, 使模型更关注有用的通道信息. 因此, 这种方法非常适用于处理较为复杂的图像.

CA 块可以被视为一个计算单元, 可以接受任何中间特征张量作为输入, 并将输出转换为与它相同大小的变换张量. CA 通过两个步骤对通道关系和远程依赖关系进行编码, 并提供精确的坐标信息嵌入和坐标注意生成操作. 图 2 显示了 CA 的结构.

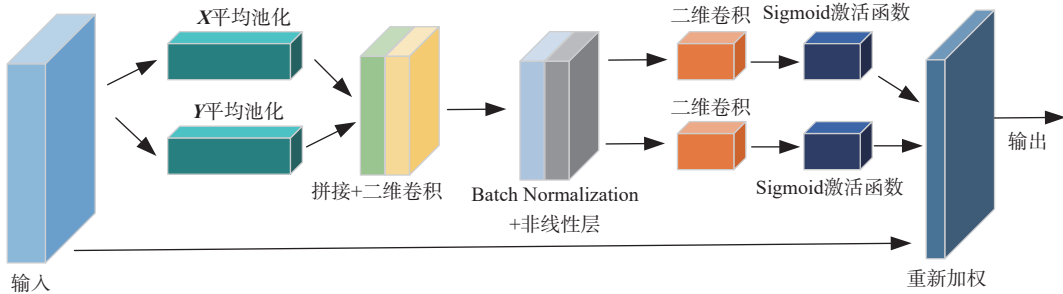


图 2 CA 结构

CA 的第一步是坐标信息嵌入. 为了更好地捕捉具有精确位置信息的远程交互, CA 将式 (3) 中的全局池化方法分解为两个平行的一维张量, 分别为式 (4) 和式 (5), 并使用大小为 $(H, 1)$ 或 $(1, W)$ 的池化核对水平和垂直坐标的每个通道进行编码.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (3)$$

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < w} x_c(h, i) \quad (4)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (5)$$

式中: $x_c(i, j)$ 为 $i \times j$ 大小的输入; z_c 为通道 c 的输出; $z_c^h(h)$ 为高度 h 的通道 c 输出; $z_c^w(w)$ 为宽度 w 的通道 c 输出.

CA 的第二步是连接式 (2) 和式 (3) 的输出, 然后使用 1×1 卷积和激活函数来降低维度并激活特征^[16]. 式 (6) 展示了这个过程. 然后, 特征被输入到 Batch Normalization 和非线性层中. 接下来, 输出 f 在空间维度上被拆分为两个独立的张量 f^h 和 f^w . 然后, F^h 和 F^w 通过两个 1×1 卷积变换转换为具有相同通道数的张量. 得到了式 (7) 和 (8).

$$f = \delta(F_1([z^h, z^w])) \quad (6)$$

$$g^h = \sigma(F_h(f^h)) \quad (7)$$

$$\mathbf{g}^w = \sigma(\mathbf{F}_w(\mathbf{f}^w)) \quad (8)$$

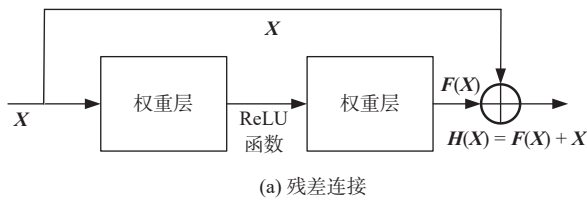
式中: $[\mathbf{z}^h, \mathbf{z}^w]$ 为特征图 \mathbf{z}^h 和特征图 \mathbf{z}^w 在空间维度上的拼接操作; δ 为非线性激活函数; \mathbf{F}_1 为 1×1 卷积; \mathbf{f} 为特征图; σ 为激活函数; $\mathbf{F}_h \in \mathbf{R}^{C/r \times H}$, $\mathbf{F}_w \in \mathbf{R}^{C/r \times W}$, \mathbf{g}^h 和 \mathbf{g}^w 为注意力权重。

最后, CA 的输出 $\mathbf{y}_c(i, j)$ 可以表示为

$$\mathbf{y}_c(i, j) = \mathbf{x}_c(i, j) \times \mathbf{g}_c^h(i) \times \mathbf{g}_c^w(j) \quad (9)$$

2.3 MixNet-s-residual

残差连接经常被用来解决 CNN 中的梯度消失和梯度爆炸等问题。残差连接指的是将 CNN 的浅层输



出和深层输出相加, 然后将相加后的数据用作下一阶段的输入。残差连接首先通过快捷连接将输入 \mathbf{X} 直接传递到输出, 然后对 \mathbf{X} 进行卷积池化等操作。将结果 $\mathbf{F}(\mathbf{X})$ 与直接传递到输出的 \mathbf{X} 相加, 得到最终的输出 $\mathbf{H}(\mathbf{X}) = \mathbf{F}(\mathbf{X}) + \mathbf{X}$ 。

提出方法构建了一种适用于 MixNet-s 的残差结构, 称为 MixNet-s-residual, 通过将残差连接与 MixNet-s 的特征结合起来。首先, \mathbf{X} 向下融入 MixNetBlock 以获得 $\mathbf{F}(\mathbf{X})$ 。然后, 将 \mathbf{X} 引入卷积层和最大池化层的串联结构以获得 $\mathbf{F}_1(\mathbf{X})$ 。最后, 将 $\mathbf{F}(\mathbf{X})$ 和 $\mathbf{F}_1(\mathbf{X})$ 连接以获得 $\mathbf{H}(\mathbf{X})$ 。图 3 展示了残差连接和 MixNet-s-residual。

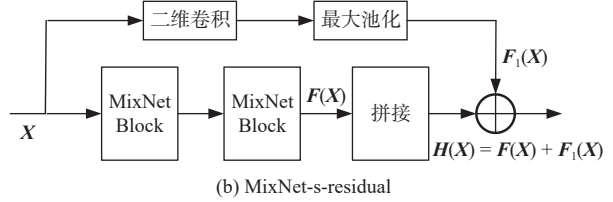
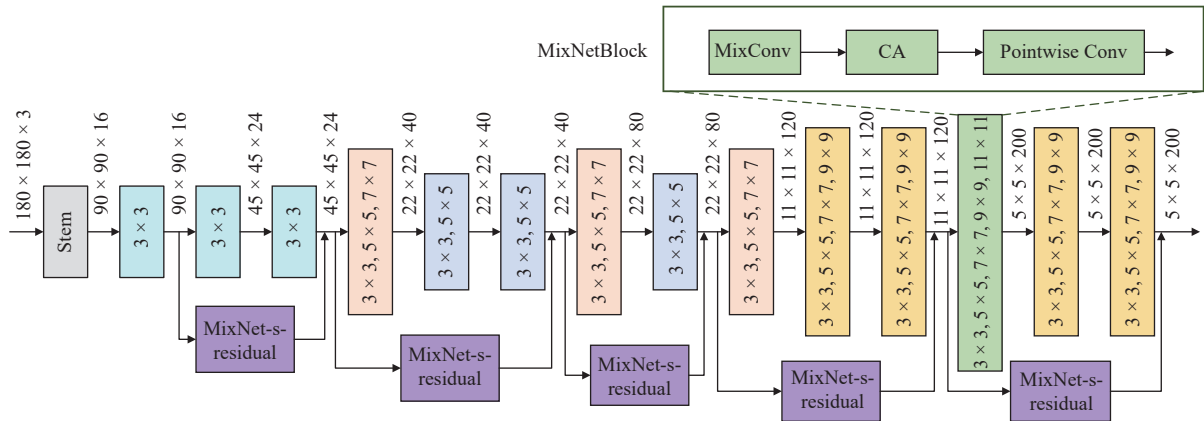


图 3 残差结构

2.4 MixNet-CA

在深度学习中, 卷积神经网络通常通过改变卷积方法、改变网络的宽度和深度以及使用注意机制等方法来提高特征提取的效果。提出方法将 MixNet-s 中的 SE 注意机制替换为 CA, 以提高网络提取特征

信息的能力。同时, 针对网络可能存在的梯度消失和爆炸问题, 将残差连接集成到网络中, 并根据网络的特征进行了改进。然后, 适当删除网络层以防止浅层网络过度学习。最终, 得到了改进的 MixNet(命名为 MixNet-CA)。MixNet-CA 的结构如图 4 所示。



3 定位系统

3.1 离线阶段

提出的定位系统包含离线阶段和在线阶段, 定位系统结构如图 5 所示。

在离线阶段, 首先在目标定位区域部署了 4 个单天线 AP。随后, 均匀分布参考点 (reference point, RP), 间隔为 0.5 m, 覆盖目标定位区域。对于每个 RP, 测量其位置处每个 AP 的 RSSI 值, 然后选择 3 个具有最

强信号强度指示 (received signal strength indication, RSSI) 值的 AP。随后, 使用 1 个带有 3 个天线的接收设备收集并规范化这 3 个 AP 的 CSI 数据, 将其缩放到 0~255。每个独立 AP 获取的 CSI 数据构成一个单通道的子图像, 随后, 这些子图像被合成为一个三通道的 RGB 图像。每个子图像本质上是一个 180×180 的实值矩阵, 其组成定义为

$$\mathbf{I}_i = [\mathbf{A}_1; \mathbf{A}_2; \mathbf{A}_3; \mathbf{P}_1; \mathbf{P}_2; \mathbf{P}_3]^T \quad (10)$$

其中: I_i 是第 i 个子图像, $i \in \{1, 2, 3\}$; A_1, A_2, A_3, P_1, P_2 和 P_3 都是大小为 180×30 的矩阵. 这些矩阵的 180 行对应于 180 个数据包, 每个矩阵有 30 列, 表示每个数据包中的 30 个子载波. 其中, A_1, A_2 和 A_3 分别是接收设备的第一、第二和第三天线接收到的 CSI 的振幅信息. P_1, P_2 和 P_3 分别是接收设备的第一和第二天线之间、第二和第三天线之间以及第三和第一天线之间的 CSI 的相位差信息. 这样构建 CSI 图像可以充分利用来自不同 AP 上不同天线的振幅和相位差信息, 产生更多样化的指纹特征.

在为所有 RP 构建 CSI 图像之后, 它们被输入到 MixNet-CA 中进行训练, 训练完成之后保存模型.

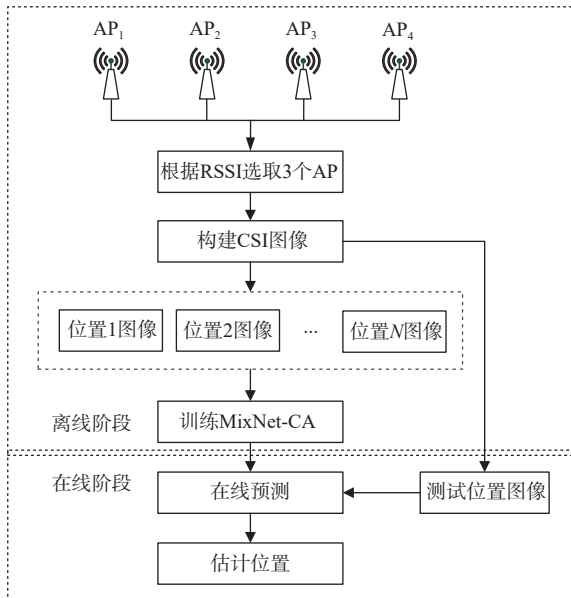


图5 定位系统结构

3.2 在线阶段

在在线阶段, 首先从目标设备收集 CSI 数据. 然后, 采用 3.1 节中描述的图像构建方法, 将 CSI 数据转换为适合输入到 MixNet-CA 模型的图像. 最后, 将构建的 CSI 图像作为输入传递给在离线阶段预训练好的 MixNet-CA 模型. 模型的输出是目标设备位于每个 RP 位置的概率值, 但实际上, 目标设备的位置可以位于定位区域的任何地方. 因此, 选择具有最高概率值的前 n 个输出, 然后使用加权质心算法估算目标设备的最终位置. 加权质心算法表示为

$$\bar{L} = \frac{\sum_{i=1}^n L_i P_i}{\sum_{i=1}^n P_i} \quad (11)$$

式中: \bar{L} 是估计的位置坐标; P_i 是第 i 个位置的概率值; L_i 是第 i 个 RP 的坐标.

4 实验结果与分析

4.1 实验配置

实验环境选择校内的一个实验室. 实验室的实际环境和平面分布分别如图 6 和图 7 所示. 实验室的尺寸为 $14 \text{ m} \times 8 \text{ m}$, 部署了 257 个 RP (图 7 中的绿圆点) 和 4 个 AP (图 7 中的红圆点). 实验室包含各种物体, 如桌椅、实验室设备和储物柜等, 导致强烈的多径效应, 使其成为非视距 (non-line-sight, NLoS) 环境. 所有相邻 RP 之间的距离为 0.5 m . 使用了 4 个单天线的 TP-Link 路由器作为发射器, 以及一台戴尔笔记本作为接收设备. 笔记本配备有 Intel 5300 无线网卡, 并运行 Ubuntu 22.04.3 LTS 操作系统. 在实验期间, 笔记本设置为监视模式, 频段和带宽分别设置为 5 GHz 和 20 MHz . 接收到的数据包之间的间隔设置为 1 ms , 所有天线与地面之间的距离为 1 m .

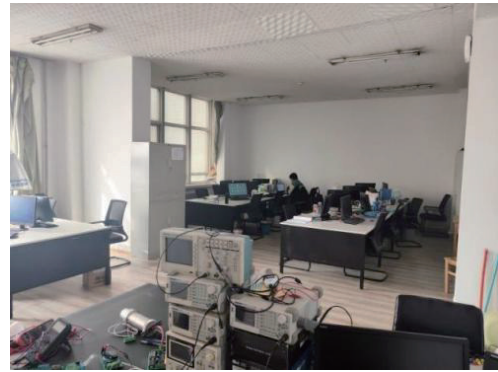


图6 实验室实际环境

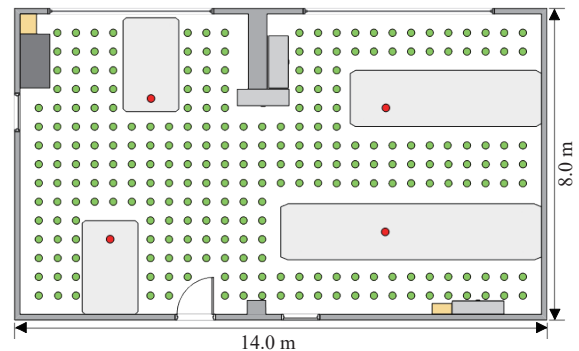


图7 实验室平面分布

对于每个 RP, 在 10 个不同的日期里, 根据 RSSI 值从 3 个最强的 AP 收集 3 s 的 CSI 数据. 每个 RP 收集了 30 000 个数据包, 前 9 天的数据用作训练数据集, 而最后一天的数据用作测试数据集. 在不同日期收集数据可以捕捉环境变化对信道的影响, 提供对信道特性更准确的描述. 训练数据集使用了 3.1 节中描述的方法转换为图像, 然后用于训练 MixNet-CA

模型. 测试数据集没有使用所有的数据包; 相反, 每个 RP 选择一部分数据包来构建测试图像. 然后将这些测试图像输入到预训练模型中以估计位置. 通过比较估计位置和实际位置来评估定位精度.

4.2 定位方法的性能比较

为了评估提出的定位方法, 采用了平均定位误差和误差的累积分布函数 (cumulative distribution function, CDF) 作为综合评价标准. 将提出方法与 CiFi、文献 [11] 的方法 (命名为 ShuffleNetV2Fi) 和文献 [12] 的方法 (命名为 CNNFi) 进行比较. 为确保实验公平, 所有方法都使用了本文收集的 CSI 数据集. 不同定位方法在实验室环境的误差累积分布如图 8 所示. 不同定位方法在实验室环境的平均定位误差、标准差和执行时间如表 2 所示.

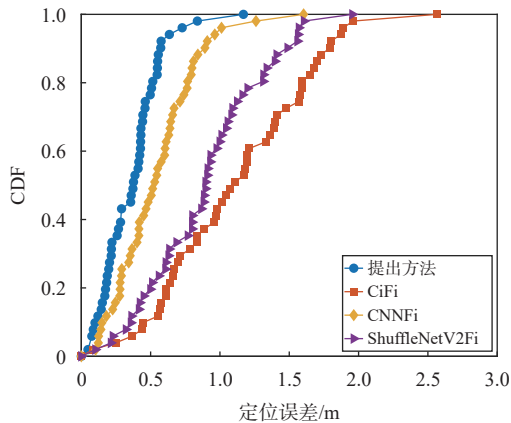


图 8 不同定位方法的累积分布函数

表 2 不同定位方法的性能比较

方法	平均误差/m	标准差/m	执行时间/s
提出方法	0.362 0	0.217 8	0.35
CiFi	1.109 0	0.541 1	0.64
CNNFi	0.537 7	0.307 4	0.52
ShuffleNetV2Fi	0.887 3	0.430 6	0.48

由图 8 和表 2 可知, 在实验室环境中, 提出方法实现了 0.362 0 m 的平均定位误差, 0.217 8 m 的标准差, 0.35 s 的执行时间, 约 80% 的测试点小于 0.5 m 的平均误差. 与 CiFi、CNNFi 和 ShuffleNetV2Fi 相比, 定位精度分别提升了约 67%、33% 和 60%, 定位执行时间分别减少了约 0.29 s、0.17 s 和 0.13 s.

4.3 不同 AP 数量对定位性能的影响

为了评估不同 AP 数量对定位性能的影响, 比较了使用不同数量的 AP 进行定位的结果. 在使用单个 AP 进行定位时, 通过构建单通道的 CSI 图像进行模

型训练. 在使用 3 个 AP 进行定位时, 通过构建三通道的 CSI 图像进行模型训练. 在使用 4 个 AP 进行定位时, 选择 3 个 RSSI 最高的 AP, 然后构建三通道的图像进行模型训练. 使用不同 AP 数量进行定位的误差累积分布如图 9 所示. 使用不同 AP 数量进行定位的平均定位误差、标准差和执行时间如表 3 所示.

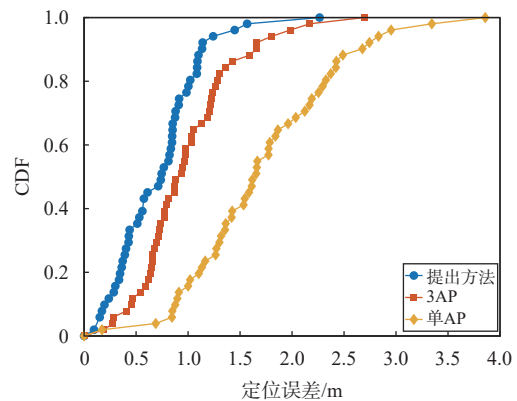


图 9 不同 AP 数量的累积分布函数

表 3 不同 AP 数量的性能比较

方法	平均误差/m	标准差/m	执行时间/s
提出方法	0.362 0	0.217 8	0.35
3AP	0.423 4	0.323 7	0.35
单AP	0.559 3	0.389 5	0.29

由图 9 和表 3 可知, 相对于仅使用一个 AP 进行定位, 使用 4 个 AP 和 3 个 AP 时的定位精度分别提高了约 35% 和 24%. 这是因为不同 AP 的信号路径和多径效应存在差异, 使得 CSI 指纹更加丰富. 此外, 使用 4 个 AP 相较于使用 3 个 AP 进行定位的精度提高了约 15%, 验证了在定位中排除最弱 RSSI 的 AP 的有效性.

4.4 训练图像数量对定位性能的影响

为了评估不同数量的训练图像对定位性能的影响, 比较了使用不同数量 (20、40、60、80、100、120) 的图像进行训练的定位结果. 不同训练图像数量的平均定位误差如图 10 所示.

由图 10 可知, 训练图像数量的增加导致定位误差持续减小. 当训练图像数量设置为 20 时, 平均误差为 1.576 6 m. 然而, 将训练图像数量提高到 80 时, 平均误差在实验室中显著减小至 0.362 m. 值得注意的是, 将训练图像数量进一步增加到 100 和 120 时, 与 80 相比, 定位误差的降低幅度较小. 因此, 在训练时间和定位精度之间取得平衡的最佳选择是每个 RP 的训练图像数量为 80.

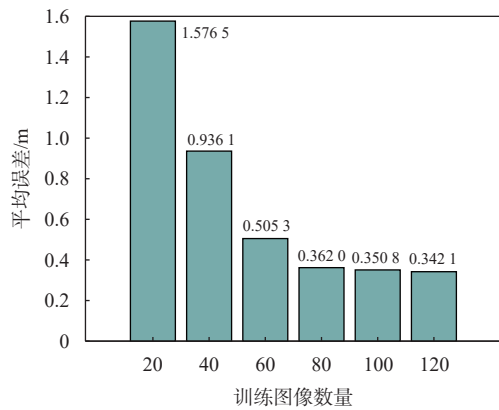


图 10 不同训练图像数量的平均误差

4.5 学习率对定位性能的影响

为了评估不同学习率对定位性能的影响,比较了使用各种学习率(0.01、0.03、0.05、0.07、0.09)进行定位的结果.不同学习率的平均误差如图 11 所示.

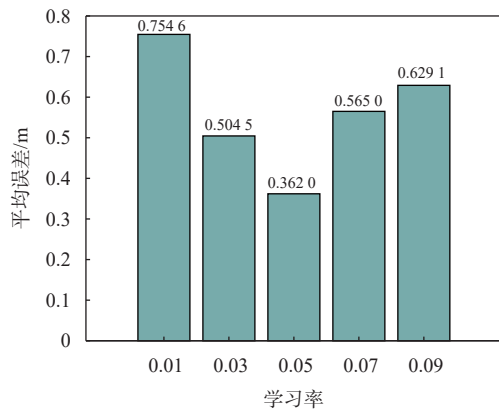


图 11 不同学习率的平均误差

由图 11 可知,较低的学习率可能妨碍 CNN 模型的收敛,而较高的学习率可能阻止其达到最佳的收敛点.考虑到这些因素,在实验中选择了学习率为 0.05 的 CNN 模型,因为它实现了最佳的定位性能.

5 结束语

提出了一种改进 MixNet 的 CSI 图像指纹室内定位方法,通过基于 RSSI 值选择 3 个 AP,减轻了信号质量差的 AP 对定位结果的影响.每个 RP 位置通过商用 Intel 5300 网卡收集多个 AP 的 CSI 数据,并组合起来构建一个三通道的 CSI 图像,这个图像捕捉了不同的信号传播路径和多径效应,提供了更丰富的指纹信息.定位系统的定位模型 MixNet-CA 对 MixNet-s 进行了改进,弃用了 SE 注意机制,转而采用 CA,以增强网络对特征信息的提取能力.为应对梯度消失和爆炸问题,巧妙地引入了残差连接,并根据网络特征进行了全面优化.此外,为避免浅层网络的过度学习,

进行了适度的网络层删除.实验结果表明,与 CiFi、CNNFi 和 ShuffleNetV2Fi 相比,提出方法在实验环境中实现了更低的平均定位误差并且拥有更快的执行速度.

参考文献

- [1] 王志恒,徐彦彦.室内定位隐私保护综述[J].通信学报,2023,44(9):188-204.
- [2] XU S H, CHEN R Z, GUO G Y, et al. Bluetooth, floor-plan, and microelectromechanical systems-assisted wide-area audio indoor localization system: apply to smartphones[J]. *IEEE transactions on industrial electronics*, 2021, 69(11): 11744-11754. DOI: 10.1109/TIE.2021.3111561
- [3] 李玉柏,孙迅.基于迁移学习提高 WiFi 室内定位信道状态信息指纹库的鲁棒性[J].电子与信息学报,2023,45(10):3657-3666.
- [4] 牟平,凌铭,胡锐.基于改进 AP 选择的融合随机森林室内定位算法[J].全球定位系统,2021,46(5):33-38.
- [5] 杨小龙,李欣玥,周牧,等.基于多维模糊映射 AP 优化的 WLAN 室内定位方法[J].电子学报,2022,50(8):1875-1884.
- [6] 俞佳豪,余敏.一种基于智能手机四向 RSS 指纹的室内定位方法[J].全球定位系统,2021,46(5):48-54.
- [7] 赵增华,童跃凡,崔佳洋.基于域自适应的 Wi-Fi 指纹设备无关室内定位模型[J].通信学报,2022,43(4):143-153.
- [8] CHEN H, ZHANG Y F, LI W, et al. ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information[J]. *IEEE access*, 2017(5): 18066-18074. DOI: 10.1109/ACCESS.2017.2749516
- [9] ZHU X Q, QU W Y, ZHOU X B, et al. Intelligent fingerprint-based localization scheme using CSI images for internet of things[J]. *IEEE transactions on network science and engineering*, 2022, 9(4): 2378-2391. DOI: 10.1109/TNSE.2022.3163358
- [10] WANG X Y, WANG X Y, MAO S W. Deep convolutional neural networks for indoor localization with CSI images[J]. *IEEE transactions on network science and engineering*, 2018, 7(1): 316-327. DOI: 10.1109/TNSE.2018.2871165
- [11] LI S H, ZENG X S, LI Y Z, et al. Convolutional neural networks based indoor Wi-Fi localization with a novel kind of CSI images[J]. *China communications*, 2019, 16(9): 250-260. DOI: 10.23919/JCC.2019.09.019
- [12] 刘帅,王旭东,吴楠.一种基于卷积神经网络的 CSI 指纹室内定位方法[J].工程科学学报,2021,43(11):1512-1521.
- [13] ZHANG F, WANG G C. Multi-branch selection fusion fine-grained classification algorithm based on coordinate attention localization[C]//2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI), 2022: 105-111. DOI: 10.1109/ICTAI56018.2022.00024

- [14] 王子辰, 陈晓艳, 王倩, 等. 基于残差自注意力连接的深度电学层析成像方法 [J]. 仪器仪表学报, 2023, 44(5): 288-301.
- [15] TAN M X, LE Q V. Mixconv: mixed depthwise convolutional kernels[J]. arXiv preprint, 2019: 1-11. DOI: [10.48550/arXiv.1907.09595](https://doi.org/10.48550/arXiv.1907.09595)
- [16] CHENG S L, WANG L J, DU A Y. Asymmetric coordinate attention spectral-spatial feature fusion network for hyperspectral image classification[J]. *Scientific reports*, 2021, 11(1): 17408. DOI: [10.1038/s41598-021-97029-5](https://doi.org/10.1038/s41598-021-97029-5)

作者简介

龙良 (1999—), 男, 硕士研究生, 研究方向为室内定位. E-mail: 2576745563@qq.com

王小鹏 (1969—), 男, 博士, 教授, 研究方向为信号处理. E-mail: wangxp1969@sina.com

李岗 (1999—), 男, 硕士研究生, 研究方向为嵌入式系统设计. E-mail: gang121815@qq.com

王江 (2001—), 男, 硕士研究生, 研究方向为物联网通信. E-mail: 2805292066@qq.com

Improved MixNet for indoor localization using CSI image fingerprints

LONG Liang, WANG Xiaopeng, LI Gang, WANG Jiang

(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: To enhance the performance of indoor localization using channel state information (CSI) fingerprints, an CSI image-based indoor localization method based on the improved MixNet model is proposed. In the offline phase, the method involves selecting the three access points (APs) with the highest received signal strength indication (RSSI) at the reference point (RP), extracting their CSI data, and converting it into image. Subsequently, the improved MixNet model is employed to train on these images and save the model. The improved MixNet model introduces coordinate attention (CA) and residual connections. Specifically, it replaces the squeeze-and-excitation (SE) attention in MixNet-s with CA to enhance the network's information representation capability and extract CSI image fingerprint features more accurately. Moreover, it incorporates residual connections, tailored to the characteristics of the MixNet-s model, to enhance the network's representation capacity and prevent overfitting. Finally, the network depth is reduced to ensure that all network layers are adequately trained. During the online phase, CSI data from the target device is collected and converted into image, and then input into the pre-trained improved MixNet model (named MixNet-CA). The final device position is estimated using a weighted centroid algorithm based on the model's output probabilities. The proposed method is validated in an indoor environment and achieve an average positioning error of 0.362 0 m.

Keywords: MixNet; coordinate attention; Wi-Fi fingerprint indoor positioning; CSI image; residual connections