

基于域自适应的 Wi-Fi 指纹设备无关室内定位模型

赵增华, 童跃凡, 崔佳洋

(天津大学智能与计算学部, 天津 300350)

摘要: 基于 Wi-Fi 指纹定位方法在大规模实际应用中存在设备多样性问题, 定位精度受到极大影响。提出了一种设备无关的 Wi-Fi 指纹室内定位模型 DeviceTransfer。该模型基于深度学习的域自适应理论, 把智能手机的设备类型作为域, 通过对抗训练来提取任务相关而设备无关的 Wi-Fi 数据特征, 并把学习到的源域位置信息迁移到目标域上。采用预训练和联合训练来提高模型训练的稳定性并加快收敛。在教学楼和商场 2 个真实场景中, 使用 4 台不同型号的智能手机验证模型的性能。实验结果表明, DeviceTransfer 能够有效提取设备无关的 Wi-Fi 数据特征。只使用一台手机在参考点采集 Wi-Fi 指纹, 使用其他型号手机在线定位也能获得较高的定位精度, 降低了定位成本。

关键词: 设备多样性; Wi-Fi 指纹定位; 室内定位; 域自适应; 深度学习

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2022069

Device-independent Wi-Fi fingerprinting indoor localization model based on domain adaptation

ZHAO Zenghua, TONG Yuefan, CUI Jiayang

College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

Abstract: In real-world large-scale deployments of indoor localization, Wi-Fi fingerprinting approaches suffer from device diversity problem which impacts the localization accuracy significantly. A device-independent Wi-Fi fingerprint indoor localization model DeviceTransfer was proposed. Based on the domain adaptation theory of deep learning, the device type of the smartphone was taken as the domain, the task-related and device-independent Wi-Fi data features were extracted through adversarial training, and the learned source domain location information was transferred to the target domain. Pre-training and joint training were employed to improve model training stability and to accelerate convergence. The performance of DeviceTransfer was evaluated using four types of smartphones in two real-world indoor environments: a school building and a shopping mall. The experimental results show that DeviceTransfer effectively extracts device-independent Wi-Fi fingerprint features. Using only one type of phone to collect Wi-Fi fingerprints, online localization using other types still achieves high localization accuracy, thus reducing localization cost significantly.

Keywords: device diversity, Wi-Fi fingerprinting localization, indoor localization, domain adaptation, deep learning

0 引言

近年来, 物联网产业蓬勃发展, 基于位置的服务引起了广泛关注。随着无线局域网的推广应用, 成千上万的无线访问接入点 (AP, access point) 部

署在室内环境中, 为室内位置服务的研究与发展奠定了基础。由于 Wi-Fi 的基础设施成本和用户设备成本都比较低, 基于接收信号强度 (RSS, received signal strength) 的 Wi-Fi 定位方法成为室内定位研究的热点。基于 RSS 的室内定位方法主要分为基于

收稿日期: 2021-12-28; 修回日期: 2022-03-13

基金项目: 国家自然科学基金资助项目 (No.61972283)

Foundation Item: The National Natural Science Foundation of China (No.61972283)

无线电传播模型的方法和基于指纹的方法^[1]。其中, 基于 RSS 的指纹定位方法由于定位精度高、部署成本低, 受到研究人员的广泛关注。该方法通常分为 2 个阶段: 离线阶段和在线阶段。离线阶段采集指纹数据(即设备接收到的周围 AP 的 RSS 值和设备所在的位置信息)构建指纹库; 在线阶段则把实时采集的 RSS 值与指纹库进行匹配, 估计出位置^[2]。

在现有研究中, 通常假设离线阶段和在线阶段都使用同一设备来收集 RSS。然而, 在实际应用中, 用户使用的设备与构建指纹库的设备型号通常不同。研究表明, 由于硬件设计、驱动程序等的差异, 不同型号的传感器具有不同的信号感测能力^[3]。即使是同一型号的传感器也会存在不同的出厂偏差。因此不同设备采集的 RSS 通常存在差异。如果忽视设备间的差异, 将导致模型的定位精度大大降低^[3-4]。这种由于设备型号、种类不同而引起的定位精度下降问题称为“设备多样性”问题。由于设备种类繁多, 为每种设备建立一个指纹库的方法显然不可行。因此, 如何只使用一种设备采集指纹库, 降低设备多样性对定位精度的影响是亟待研究的重要课题。

为此, 一些研究通过设备校准来减小不同设备观测值之间的差异^[5-6]。例如, Tsui 等^[5]通过收集不同设备的 RSS 数据来训练转换模型, 补偿训练设备和测试设备之间的 RSS 差值。然而, 该方法需要收集不同设备在同一地点观测到的 RSS 值以获得训练数据, 消耗大量的人力物力。为此, Huang 等^[6]提出基于直方图的 RSS 校准方法, 将一对一的 RSS 约束放宽到分布之间的约束。此外, 还有一些研究通过使用线性回归的方法来挖掘不同设备采集的 RSS 值之间的差异^[1,7-8], 将多个设备的指纹映射到单个指纹库, 以此来消除设备多样性带来的影响。这种方法虽然不需要收集成对的 RSS 值, 但要求在训练阶段使用多个设备的带有位置信息的 RSS 数据, 数据采集成本依然很高。

为了降低数据采集成本, 消除设备多样性对定位精度的影响, 本文提出了一种设备无关的 Wi-Fi 指纹室内定位模型 DeviceTransfer。该模型基于域自适应深度学习理论^[9], 把用于采集指纹的设备类型作为源域, 待定位的设备类型作为目标域, 只需要使用一种设备(源域)采集指纹, 通过域自适应就可以把源域的位置信息迁移到待定位的其他类型设备(目标域)采集的 RSS 数据上。由于不需要对

目标域设备采集的 RSS 数据进行位置标定, 因此可以很方便地采集待定位设备的数据, 大大降低了数据采集成本。

DeviceTransfer 由特征提取器、解码器、生成器、鉴别器和定位器组成。特征提取器用于提取来自不同类型设备的 RSS 数据的设备无关特征, 并且使这些特征在定位时能够得到较高的定位精度。为此, 特征提取器与解码器结合, 构成自动编码器, 学习每个类型设备的数据特征; 与生成器和鉴别器结合, 通过对抗性训练学习与设备类型无关的数据特征; 与定位器结合, 学习与定位任务相关的特征。DeviceTransfer 把设备类型作为条件输入生成器和鉴别器, 因此只需要使用一组生成器和鉴别器, 提高了模型处理多种设备类别的能力。为了能够处理离散的 RSS 数据, DeviceTransfer 采用了卷积神经网络(CNN, convolutional neural network)结构来捕获数据特征。此外, 还设计了预训练策略以提高模型训练的稳定性并加快收敛。

本文使用 4 台不同型号的智能手机在教学楼和商场 2 个真实的室内场景中采集了 Wi-Fi 数据集, 对 DeviceTransfer 进行了性能评价, 并与其他同类算法和模型进行了对比。结果表明, DeviceTransfer 能够提取 RSS 数据的设备无关特征, 使用一种设备采集 Wi-Fi 指纹, 其他设备也可以取得较高的定位精度。DeviceTransfer 有效地消除了设备多样性对定位精度的影响, 定位性能优于其他同类算法和模型。

1 相关工作

1.1 室内定位设备多样性研究

由于不同设备采集的 RSS 值通常存在差异, 若直接使用离线设备采集的数据建立定位模型用于在线设备, 会导致模型定位精度大大降低, 因此不能忽视这种差异。为了降低设备多样性对模型定位精度的影响, 现有的大多数方法都着力于缩小不同设备采集的 RSS 值的差异。一种解决方案依靠 RSS 校准来解决问题。文献[5]利用转换模型来缩小训练设备和测试设备之间的 RSS 差异, 要求使用 2 个设备在同一地点收集 RSS 值, 组成成对 RSS 训练数据。然而, 用户设备的型号无法提前得知, 因此很难自适应地选择合适的模型, 并且该方案对每个新的设备需要训练新的转换模型, 不具备可扩展性。文献[6]使用基于直方图的 RSS 校准方法, 将一对一的 RSS 约束放宽到分布的约束。该方法既不需要

手动配对收集数据，也不需要预知设备信息。一旦确定校准模型，所有新来的 RSS 都可以被校准。还有一些方案使用线性回归的方法。假设不同设备之间的 RSS 值具有线性关系，通过线性回归方法，将不同设备采集到的 RSS 值映射到相同的信号空间，构建一个统一的指纹库^[1,7-8]。然而这种方法需要收集不同设备的大量 RSS 值，并且要求标记所有的 RSS 值，因此数据采集成本极大。此外，相对特征也可以用于消除设备类型的影响，如使用 RSS 差值、RSS 比率和 RSS 顺序等进行定位。由于相对特征对设备不那么敏感，因此可以减少设备多样性的影响，而不需要 RSS 校准步骤^[10]。然而相对 RSS 特征在相邻区域往往是相似的，因此这种方法是以降低精度为代价的。近年来，随着深度学习的普及，有些研究使用卷积神经网络来解决设备多样性问题^[11-12]。虽然这些研究不属于室内定位领域，但是给了本文很多启发。

1.2 域自适应模型

域自适应模型是迁移学习的一种，适用于数据标记成本高或者标记困难，而不需要标记的数据获取便捷的应用场景^[12]。有标记的数据为源域，无标记的数据为目标域^[13]。假定源域和目标域的数据分布不同，但是任务相同。域自适应模型需要学习源域和目标域的域无关特征，并将源域的标记信息迁移到目标域上。

根据域无关特征提取的方法，域自适应模型可以分为最小化源域和目标域分布的距离散度的方法^[14-15]、基于重构的方法^[16-17]和基于对抗的方法^[18-21]。基于对抗的方法通过对抗训练来实现域自适应，域对抗神经网络（DANN, domain-adversarial neural network）^[19]就是这类模型的代表。DANN 结合对抗网络框架与深度学习技术，使用生成器生成来自不同域的数据，用鉴别器判别输入的样本是来自真实的数据还是生成器。通过博弈不断提高生成器生成数据的精度，获取域无关特征。

MotionTranformer^[22]成功地将域自适应模型应用到基于惯性运动单元（IMU, inertial motion unit）序列的轨迹追踪中，只需要标定一种姿态的 IMU 序列，即可高精度地追踪其他姿态的 IMU 序列。DeviceTransfer 借鉴了 MotionTranformer 的域自适应框架，不同的是，DeviceTransfer 把设备类型作为条件输入生成器和鉴别器，因此只需要使用一个生成器和一个鉴别器，而不是每个设备类型使用

一个，提高了模型处理多种设备类别的能力。为了能够处理离散的 RSS 数据，DeviceTransfer 使用了不同的网络结构来捕获数据特征。

2 室内定位设备多样性问题及其对定位精度的影响

2.1 室内定位设备多样性问题

在实际应用中，不同用户使用的移动设备通常种类和型号繁多。而不同类型设备的 Wi-Fi 信号接收器件感测能力不同，因此接收的信号强度指示（RSSI, received signal strength indicator）值通常存在较大差异。为了验证设备多样性问题，使用 4 台不同型号的智能手机（vivo X5M、Redmi note 1S、HUAWEI Mate8 和 HUAWEI Mate9），在学校某教学楼的同一位置同时收集 Wi-Fi RSSI 数据。

图 1 展示了这 4 台手机接收到的同一个 AP 的 RSSI 值。从图 1 可以看到，即使在同一位置，同一时间段内 4 台设备接收到的同一 AP 的 RSSI 都会呈现出较明显的差异。因此使用 Wi-Fi 指纹定位时，设备多样性问题是实际存在的。

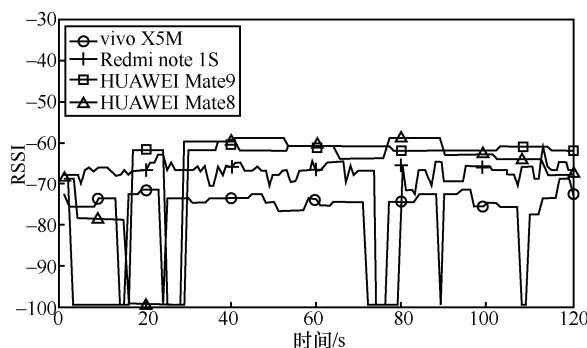


图 1 4 台不同型号的手机在同一位置、相同时间段内接收到同一 AP 的 RSSI 数据

2.2 设备多样性对定位精度的影响

为了进一步验证设备多样性对定位精度的影响，在教学楼进行了定位实验，定位算法选用深度神经网络（DNN, deep neural network）模型（网络结构和 3.2 节中的定位器相同）。用 4 台智能手机分别采集 Wi-Fi 指纹库。使用其中一台手机的指纹库训练定位模型，使用其余手机采集的 RSSI 数据进行定位测试。

实验结果如图 2 所示。其中，横轴为采集指纹库所用的手机型号，纵轴为在不同设备上测试所得到的平均定位误差。以 vivo 手机采集指纹库为例（最左侧一组），使用 vivo 定位的误差最小，

为 0.81 m。使用其他手机定位的误差显著增加，最小的定位误差为 1.24 m，最大达到 1.90 m，是 vivo 定位误差的 2 倍多。其他几组数据也有类似的结论，即使用和采集指纹库相同的手机定位精度高，而使用其他型号的手机定位精度明显降低。

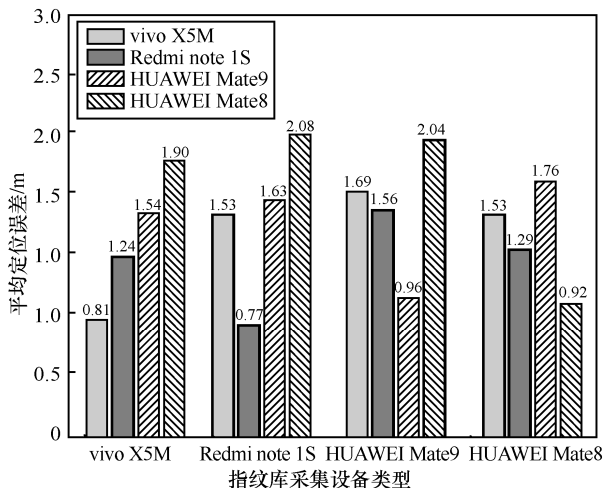


图 2 设备多样性对定位精度的影响

这说明设备多样性极大地影响了定位精度。尤其在实际应用中，用户使用的智能手机型号众多，而且新的手机型号层出不穷，为每种型号的手机建立一个指纹库成本极高，也不现实。因此，在基于 Wi-Fi 指纹的定位算法中，亟须消除设备多样性的影响，以提升用户的室内定位精度。

3 DeviceTransfer 模型

3.1 模型概述

DeviceTransfer 模型旨在通过深度域自适应来提取 Wi-Fi 数据的设备无关特征表示，只需要采集一种设备的 Wi-Fi 指纹库，使用其他设备定位也能获得较高的定位精度。在降低指纹库采集成本的同时，充分消除设备多样性的影响。

DeviceTransfer 模型的输入为在待定位区域接收到的 Wi-Fi RSSI 数据 $x = (rssi_1, rssi_2, \dots, rssi_n)$ ，其中 n 为该区域能够接收到信号的 AP 数量。把设备的类型作为域，记为 d 。 $d = \{s, t\}$ ，其中 s 表示源域， t 表示目标域。源域只有一种设备类型，目标域可以有多个， $t = t_1, t_2, \dots, t_k$ ， k 为目标域设备类型的数量。源域数据是由源域设备在参考点采集的 Wi-Fi 指纹，记为 x^s ，带有位置标签 $y^s \in R^2$ 。目标域数据则不带有位置标签，记为 x^t ，可以由用户手持目标域设备在待定位区域随意行走采集。由于目标域数据不

需要专业人员在固定的参考点采集，因此可以很方便地获取到，采集成本相较于源域数据大大降低。

DeviceTransfer 模型框架如图 3 所示，主要包括特征提取器、解码器、生成器、鉴别器和定位器 5 个模块。特征提取器通过与其他模块协作来提取 Wi-Fi RSSI 数据 x 的设备无关、任务（定位）相关的特征。具体来讲，特征提取器与解码器组成一个条件自动编码器，将输入的 Wi-Fi RSSI 数据 x 根据条件“设备类型” d 编码到潜在特征空间，再从特征空间中恢复出原始数据。通过每种类型自动编解码来学习各个设备类型的数据特征。特征提取器与生成器和鉴别器协作，构成条件生成对抗网络。通过对抗训练学习与设备无关的特征，特征提取器还和定位器协作提取数据的定位任务相关特征。在模型训练过程中，通过梯度下降，定位器激励特征提取器提取出与定位任务相关的特征以进行准确的位置估计。

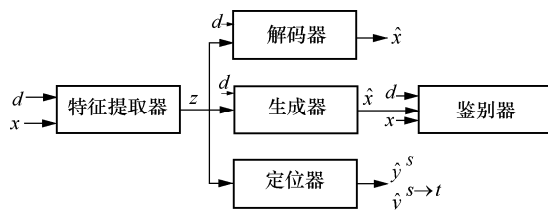


图 3 DeviceTransfer 模型框架

3.2 模型设计

特征提取器。特征提取器负责提取输入数据的特征，用函数表示为 $f_{enc}(\cdot)$ 。对输入的 RSSI 数据 x 和域（设备类型） d 进行条件编码，得到潜在空间的数据特征 z 为

$$z = f_{enc}(x, d) \tag{1}$$

解码器。条件编码器和条件解码器共同构成一个条件自动编码器 (CVAE, conditional variational autoencoder) [23]，根据输入的条件域 d 从特征空间重建 RSSI 数据，以此来学习不同类型设备的数据特征。令 $f_{dec}(\cdot)$ 表示解码器的函数， \hat{x}^s 和 \hat{x}^t 分别表示从 x^s 和 x^t 重建的数据，则有

$$\begin{aligned} \hat{x}^s &= f_{dec}(z, s) = f_{dec}(f_{enc}(x^s, s), s) \\ \hat{x}^t &= f_{dec}(z, s) = f_{dec}(f_{enc}(x^t, t), t) \end{aligned} \tag{2}$$

条件自动编码器的损失函数是以设备类型为条件重建 x 的重建损失，记为 \mathcal{L}_{rec} 。使用均方误差 (MSE, mean square error)，则 \mathcal{L}_{rec} 可以表示为

$$\mathcal{L}_{\text{rec}} = \mathbb{E}_{x^s \sim p(x^s), x^t \sim p(x^t)} [\|\hat{x}^s - x^s\|_2 + \|\hat{x}^t - x^t\|_2] \quad (3)$$

通过最小化 \mathcal{L}_{rec} 损失，把设备类型 d 的 RSSI 数据 x^d 编码到潜在特征空间并恢复出 x^d ，鼓励特征提取器学习各种类型设备的数据特征。

生成器与鉴别器。生成器与鉴别器和特征提取器共同构成条件生成对抗网络 (CGAN, conditional generative adversarial network)，以设备类型 d 为条件。和原始 CGAN 不同，为了增加模型的稳定性^[24]，把特征提取器的输出作为生成器的输入，而不是使用随机变量^[25]。特征提取器将真实的 RSSI 数据编码到潜在特征空间，然后生成器从潜在特征空间中生成某设备类型 d 的 RSSI 数据 (假数据)。

为了使特征提取器学习到和设备类型无关的 RSSI 数据特征，源域的 RSSI 数据 x^s 生成到目标域后不能让鉴别器识别出来，反之亦然。用 $\hat{x}^{s \rightarrow t}$ 表示从源域数据生成的目标域数据，则这一数据转换操作 $G_{s \rightarrow t}(x^s, t)$ 可以表示为

$$\hat{x}^{s \rightarrow t} = G_{s \rightarrow t}(x^s, t) = f_{\text{gen}}(f_{\text{enc}}(x^s, s), t) \quad (4)$$

同理，从目标域数据转换成源域的数据 $G_{t \rightarrow s}(x^t, s)$ 可以表示为

$$\hat{x}^{t \rightarrow s} = G_{t \rightarrow s}(x^t, s) = f_{\text{gen}}(f_{\text{enc}}(x^t, t), s) \quad (5)$$

把条件生成对抗网络的损失称作 GAN 损失，记为 \mathcal{L}_{GAN} 。GAN 损失包括两部分数据的损失： $\mathcal{L}_{\text{GAN}}^{s \rightarrow t}$ 和 $\mathcal{L}_{\text{GAN}}^{t \rightarrow s}$ 。 $\mathcal{L}_{\text{GAN}}^{t \rightarrow s}$ 用于鉴别源域的真实数据和从目标域生成的假数据。 $\mathcal{L}_{\text{GAN}}^{s \rightarrow t}$ 则用于鉴别目标域的真实数据和从源域生成的假数据。令 D 表示鉴别器的函数，则损失函数 \mathcal{L}_{GAN} 可以表示为

$$\begin{aligned} \mathcal{L}_{\text{GAN}} &= \mathcal{L}_{\text{GAN}}^{t \rightarrow s} + \mathcal{L}_{\text{GAN}}^{s \rightarrow t} \\ \mathcal{L}_{\text{GAN}}^{t \rightarrow s} &= \mathbb{E}_{x^s \sim p(x^s)} [(D(x^s) - 1)^2] + \mathbb{E}_{x^t \sim p(x^t)} [D(\hat{x}^{t \rightarrow s})^2] \\ \mathcal{L}_{\text{GAN}}^{s \rightarrow t} &= \mathbb{E}_{x^t \sim p(x^t)} [(D(x^t) - 1)^2] + \mathbb{E}_{x^s \sim p(x^s)} [D(\hat{x}^{s \rightarrow t})^2] \end{aligned} \quad (6)$$

通过最小化 GAN 损失，生成器和鉴别器展开博弈。生成器用从源域数据提取的特征来生成目标域的假数据，用从目标域数据提取的特征来生成源域的假数据，用假数据来欺骗鉴别器。另一方面，鉴别器努力识别出假数据和真数据。通过博弈，驱使特征提取器提取源域数据和目标域数据的设备无关特征。

定位器。为了实现设备无关的位置估计，定位器并不是直接使用原始 RSSI 数据进行定位，而是使用 RSSI 数据的设备无关特征，即原始 RSSI 数据先由特征提取器提取出设备无关特征，然后输入定位器。

由于只有源域 RSSI 数据带有位置标签而目标域数据没有标签，因此只能使用源域设备的指纹库来训练定位器。用 f_{loc} 表示定位器的函数，则定位器估计出的源域数据 x^s 的位置 \hat{y}^s 可以表示为

$$\hat{y}^s = f_{\text{loc}}(f_{\text{enc}}(x^s, s)) \quad (7)$$

另一方面，既然特征提取器能够提取与设备类型无关的数据特征，那么源域数据 x^s 经过特征提取转换到目标域后的 $\hat{x}^{s \rightarrow t}$ 也应该能对应到原来的位置上。用 $\hat{y}^{s \rightarrow t}$ 表示 $\hat{x}^{s \rightarrow t}$ 的估计位置，则有

$$\hat{y}^{s \rightarrow t} = f_{\text{loc}}(f_{\text{enc}}(\hat{x}^{s \rightarrow t}, t)) \quad (8)$$

因此，定位器的目标函数 \mathcal{L}_{loc} 设计为

$$\mathcal{L}_{\text{loc}} = \mathbb{E}_{(x^s, y^s) \sim p(x^s, y^s)} [\|y^s - \hat{y}^s\|_2 + \|y^s - \hat{y}^{s \rightarrow t}\|_2] \quad (9)$$

通过最小化定位损失，既能鼓励特征提取器提取与定位任务相关的特征，也能加强对设备无关特征的提取能力。

DeviceTransfer 模型训练好后，可以使用训练好的特征提取器和定位器来估计目标域设备数据 x^t 的位置 \hat{y}^t ，表示为

$$\hat{y}^t = f_{\text{loc}}(f_{\text{enc}}(x^t, t)) \quad (10)$$

样本级循环一致性约束。根据文献[26]，采用循环一致性正则化的方法可以增强特征提取器域无关特征的能力。因此，本文设计了样本级循环一致性约束和特征级循环一致性约束。

样本级循环一致性约束的基本思想是源域数据 x^s 转换成目标域数据 $x^{s \rightarrow t}$ 后再输入特征提取器然后转换回源域 $\hat{x}^{s \rightarrow t \rightarrow s}$ ，应该能和 x^s 一致，即

$$\hat{x}^{s \rightarrow t \rightarrow s} = G_{t \rightarrow s}(G_{s \rightarrow t}(x^s, t), s) \quad (11)$$

其中， $G_{s \rightarrow t}$ 和 $G_{t \rightarrow s}$ 的定义参见式(4)和式(5)。

样本级循环一致性损失定义为 $\mathcal{L}_{\text{xcycle}}$ ，表示为

$$\mathcal{L}_{\text{xcycle}} = \mathbb{E}_{x^s \sim p(x^s)} \|\hat{x}^{s \rightarrow t \rightarrow s} - x^s\|_2 + \mathbb{E}_{x^t \sim p(x^t)} \|\hat{x}^{t \rightarrow s \rightarrow t} - x^t\|_2 \quad (12)$$

特征级循环一致性约束。特征级循环一致性是指源域数据 x^s 和由源域生成的目标域数据转换回源域后的数据 $\hat{x}^{s \rightarrow t \rightarrow s}$ 应该在潜在特征空间中保持一致性。反之，目标域数据和循环转换到目标域的数

据也应该在潜在特征空间中保持一致。

假设 z^s 为特征提取器从源域数据提取的特征, $\hat{z}^{s \rightarrow s'}$ 为特征提取器从循环转换后的数据 $\hat{x}^{s \rightarrow s'}$ 中提取的特征, z^t 和 $\hat{z}^{t \rightarrow s'}$ 分别为目标域数据及其循环后的数据特征, 则特征级循环一致性损失 $\mathcal{L}_{\text{zycycle}}$ 表示为

$$\mathcal{L}_{\text{zycycle}} = \mathbb{E}_{x^s \sim p(x^s)} \left[\|z^s - \hat{z}^{s \rightarrow s'}\|_2 \right] + \mathbb{E}_{x^t \sim p(x^t)} \left[\|z^t - \hat{z}^{t \rightarrow s'}\|_2 \right] \quad (13)$$

通过最小化样本级和特征级循环一致性损失, 能够降低原始数据和生成数据在样本空间和特征空间的差异, 使特征提取器提取的数据特征具有更好的设备无关性。

目标函数。综上, 为模型定义的损失函数有重建损失、GAN 损失、定位损失、样本级循环一致性损失和特征级循环一致性损失。DeviceTransfer 模型的目标函数 $\mathcal{L}_{\text{total}}$ 为上述各项损失之和, 即

$$\mathcal{L}_{\text{total}} = \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} + \lambda_{\text{gan}} \mathcal{L}_{\text{GAN}} + \lambda_{\text{loc}} \mathcal{L}_{\text{loc}} + \lambda_{\text{xc}} \mathcal{L}_{\text{xcycle}} + \lambda_{\text{zc}} \mathcal{L}_{\text{zycycle}} \quad (14)$$

其中, λ_{gan} 、 λ_{rec} 、 λ_{loc} 、 λ_{xc} 、 λ_{zc} 为用于调整各部分权重的超参数。

3.3 训练策略

由于存在大量的无标签数据, 直接使用所有数据训练模型会导致模型不稳定、收敛慢。为此, 设计了预训练和联合训练。先使用有标签的源域数据(即指纹库)对模型各个部分单独进行训练, 即预训练。然后保留预训练的参数, 对所有模块进行联合训练。

预训练。将模型分成 3 个部分, 分别进行预训练: 特征提取器和解码器组成的 CVAE、特征提取器和生成器、特征提取器和定位器。对这 3 个部分进行预训练的损失函数分别为式(3)、式(6)和式(9)中的源域数据部分, 选用 Adam^[27] 优化器进行小批量梯度下降训练。

模型的预训练是一个有监督的学习过程, 通过该过程, 特征提取器可为重建和位置预测学习源域的特征表示。预训练后, 特征提取器、解码器、生成器和定位器的参数被保留下来, 参与联合训练。这可以为联合训练提供高质量的初始化参数, 带来更好的泛化性能。

联合训练。预训练之后, 对整个模型进行联合训练。其中, 定位器、特征提取器、解码器和生成器被联合起来与鉴别器进行对抗训练。本文模型中

的对抗训练通过梯度反转层(GRL, gradient reversal layer)来实现^[20]。鉴别器与特征提取器通过 GRL 连接, 在反向传播的训练过程中, GRL 将梯度乘以某个负常数, 使 GRL 前后网络训练目标相反, 以实现对抗效果。选用 Adam^[27] 优化器进行小批量梯度下降训练。一个小批量的样本数量通常设置为 2 的倍数。在一个训练周期内分批次训练所有样本。一个模型通常需要经过多个周期的训练才能收敛。联合训练由于是在预训练的基础上, 因此可以加速训练的收敛, 实现模型高效、稳定的训练。

4 实验与性能评价

4.1 实验数据

为了验证 DeviceTransfer 模型在真实场景中的性能, 本文用 4 台不同型号的智能手机在 2 个室内场景中采集 Wi-Fi RSSI 数据。按照实验场景的不同, 构建了 2 个数据集: 教学楼数据集和商场数据集。在这 2 个数据集下进行实验验证。下面详细介绍数据集的数据采集方式。

实验设备。用于采集数据的 4 台智能手机分别为 vivo X5M、Redmi note 1S、HUAWEI Mate9 和 HUAWEI Mate8, 每台手机代表一个域。DeviceTransfer 模型使用 pytorch 1.4.0 的深度学习平台^[28]实现, 部署在实验室的一台服务器上, CPU 为 Intel Core Processor I7-7700K, GPU 为 GeForce GTX 1080Ti。

实验场景。2 个室内场景分别为天津大学某教学楼走廊和某商场一层, 场景总测试面积约为 8 020 m²。室内平面结构如图 4 和图 5 所示, 参考点与测试点的位置互不重叠。场景信息如表 1 所示, 包含场景面积、AP 个数、参考点个数和测试点个数。2 个场景中的 AP 均为环境中已经部署好的, 不需要为实验专门部署。

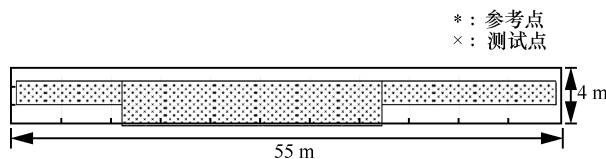


图 4 教学楼走廊数据采集点分布

数据获取方式。在教学楼走廊中, 2 个相邻参考点之间的间隔为 1.6 m, 测试点之间的间隔为 1.6 m。把 4 台手机固定在支架上, 放在参考点和测试点上采集环境中 AP 的 RSSI 数据。测试点采集时间为 8 s,

采集频率为 5 Hz。在商场中，2 个相邻参考点之间的间隔为 5 m，测试点之间的间隔为 9 m。采集人员手持 2 台智能手机在每个采集点采集数据。参考点采集时间为 60 s（采集较长时间以消除 RSSI 的波动），测试点采集时间为 5 s，采集频率为 5 Hz。

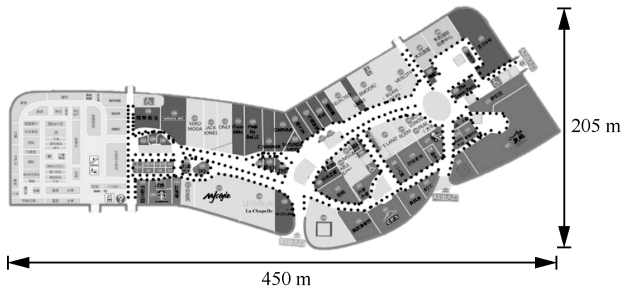


图5 商场一层平面

表1 场景信息

场景	场景面积/m ²	AP/个	参考点/个	测试点/个
教学楼走廊	220	124	97	92
商场一层	7 800	1 028	300	99

数据集。教学楼和商场采集的数据分别汇集成 2 个数据集：教学楼数据集和商场数据集。在实验过程中每个数据集划分为三部分：训练集、验证集和测试集。其中训练集和验证集的数据由参考点采集的数据得到，将数据按 9:1 的比例随机划分为训练集和验证集 2 个部分。源域的 RSSI 数据带位置标签，目标域设备的数据不带位置标签。实验中为方便起见选定 vivo 手机为源域，其他手机为目标域。测试集由测试点采集的数据构成。

在每个训练周期中，使用训练集数据训练模型，然后将验证集的数据代入模型进行计算；使用验证集的计算结果来选取每轮实验中最佳的训练周期，并保留模型参数；使用测试集进行模型性能测试。

4.2 实验方法与参数设置

实验方法。本文在教学楼和商场 2 个数据集上训练 DeviceTransfer 模型，然后进行测试。测试时，把待测位置处的 RSSI 数据输入 DeviceTransfer 的特征提取器，编码到潜在空间，得到设备无关特征；然后输入定位器估计该 RSSI 数据所在的位置。

为了验证 DeviceTransfer 的设备无关性能，设计了两组对比实验。第一组验证 DeviceTransfer 的设备无关特征提取的有效性，将实验结果和“直接定位模型”相对比。直接定位模型使用带标签的源域数据训练定位器，其网络结构和 DeviceTransfer

模型中的定位器相同。使用训练好的定位器对目标域数据进行位置估计。

第二组实验验证 DeviceTransfer 模型的设备无关的定位性能，与以下 3 种模型和算法进行对比。

1) 域自适应模型 DANN^[20]。DANN 的特征提取器和定位器的网络结构与 DeviceTransfer 模型中的相应模块保持一致。

2) 基于深度学习的定位模型 CNNLoc^[29]。CNNLoc 是最近提出的一种基于 CNN 的 Wi-Fi 指纹室内定位模型。

3) 设备无关定位算法 KNN-cosine^[30]。KNN-cosine 是在 Han 等^[30]算法的基础上设计的设备无关的 KNN 算法。该算法用余弦相似度代替指纹匹配中常用的欧氏距离，能较好地消除室内定位中的设备多样性问题。

参数设置。图 6 为 DeviceTransfer 各模块的网络结构。以特征提取器为例解释图中符号含义，如图 6(a)所示，对于输入的 Wi-Fi 指纹，首先，通过一个全连接(FC, full-connection)层将其转为 1 024 大小的向量。其次，通过 Reshape 将其整形为 32×32 的矩阵。再次，将整形的指纹输入 2 个 CNN 块(ConvBlk, convolutional block)，每个 CNN 块包含一个卷积层、批量归一化(BN, batch normalization)层^[31]、激活函数(LeakyReLU)^[32]和一个平均池化(average pooling)层。最后，经过几个全连接层输出。解码器与生成器中采用的是去卷积神经网络(DCNN, de-convolutional neural network)^[33]，它是卷积的逆过程。每个 DCNN 块(DeConvBlk, de-convolution block)中包含一个反卷积层、BN 层、激活函数 LeakyReLU 和一个上采样(upsample)层。鉴别器是一个使用了谱归一化(spectral normalization)技术的 CNN 结构^[34]。图 6(c)是定位器使用深度残差网络(ResNet, residual network)^[35]时的网络结构，其中 ResBlk-0 和 ResBlk-1 是 2 种不同的残差块。

总体而言，在教学楼数据集上，DeviceTransfer 模型中特征提取器和鉴别器使用 CNN 结构，解码器和生成器使用 DCNN 结构，定位器分别使用 DNN、CNN 和 ResNet 做对比。在商场数据集上，DeviceTransfer 模型中鉴别器使用 CNN 结构，特征提取器使用 ResNet 结构，解码器和生成器使用 DeResNet 结构，定位器分别使用 DNN、CNN 和 ResNet 做对比。

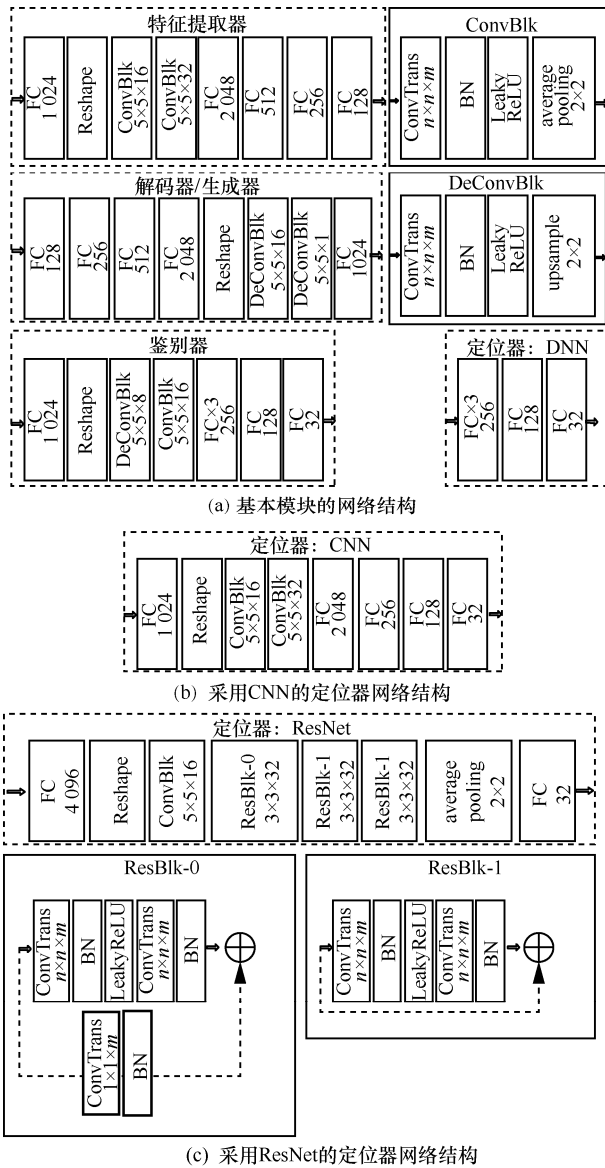


图6 DeviceTransfer 各模块的网络架构

DeviceTransfer 模型训练过程中，设置超参数 $\lambda_{rec} = 1$ 、 $\lambda_{gan} = 1$ 、 $\lambda_{loc} = 0.01$ 、 $\lambda_{xc} = 1$ 、 $\lambda_{zc} = 1$ 。优化器学习率设置为 0.000 2，小批量中样本数据数量设置为 16。预训练 50 个周期，联合训练 100 个周期。直接定位模型训练周期设置为 100。所有实验均进行 5 次，实验结果为 5 次结果的平均值。

4.3 实验结果与性能分析

1) DeviceTransfer 设备无关特征提取的有效性验证

为了验证 DeviceTransfer 模型提取设备无关特征的有效性，将 DeviceTransfer 与直接定位模型进行比较。本文选择 3 种不同的网络结构作为定位器：DNN、CNN、ResNet。DeviceTransfer 模型的定位

器结构选用以上 3 种网络结构时，分别简称为 DT-D、DT-C 和 DT-R。以 vivo X5M 手机作为源域设备，其余手机作为目标域设备，统计所有设备在不同定位模型中的定位误差。本节分别在教学楼和商场数据集上进行实验，在 2 个数据集上，不同模型的对比结果如图 7 和图 8 所示。

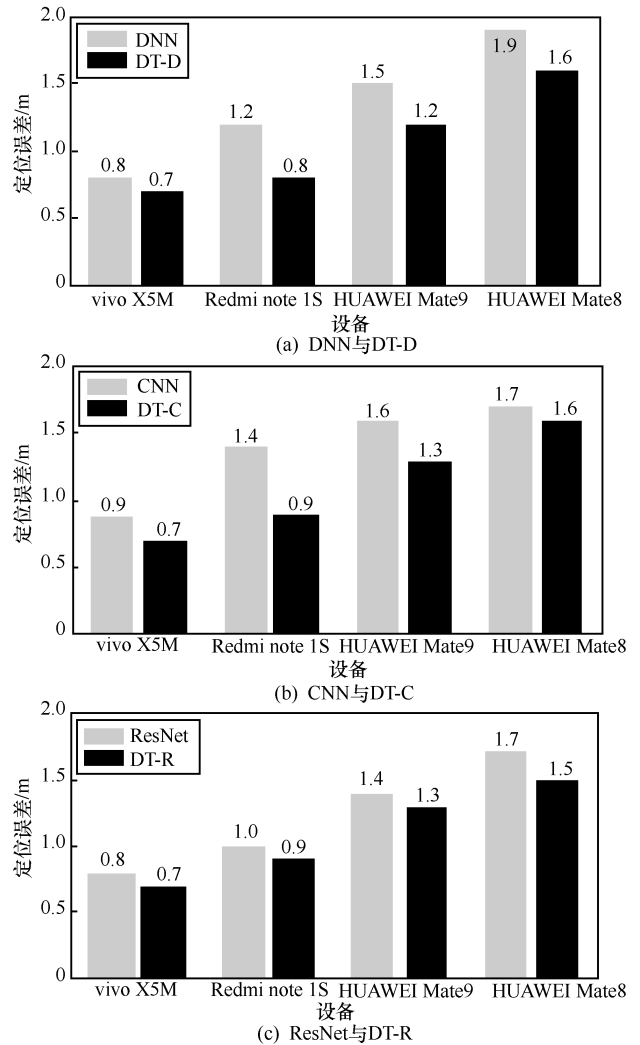


图7 教学楼数据集上 DeviceTransfer 设备无关性能的验证

不同设备数据在经过 DeviceTransfer 模型提取设备无关特征后再进行定位得到的定位误差比直接使用源域定位器的定位误差小。图 7(a)中目标域设备 Redmi note 1S 在 DNN 上定位误差为 1.2 m，在 DT-D 上定位误差为 0.8 m，域适应后定位误差比直接定位误差下降了 0.4 m，与源域设备定位误差的差距也从 0.4 m 下降到 0.1 m。目标域设备 HUAWEI Mate9 和 HUAWEI Mate8 在域适应后的定位误差也都下降了 0.3 m。图 8(a)中目标域设备在 DNN 上定位误差为 11.7 m，在 DT-D 上定位误差为

5.5 m, 域适应后定位误差比直接定位误差下降了 6.2 m, 定位精度提升效果明显。

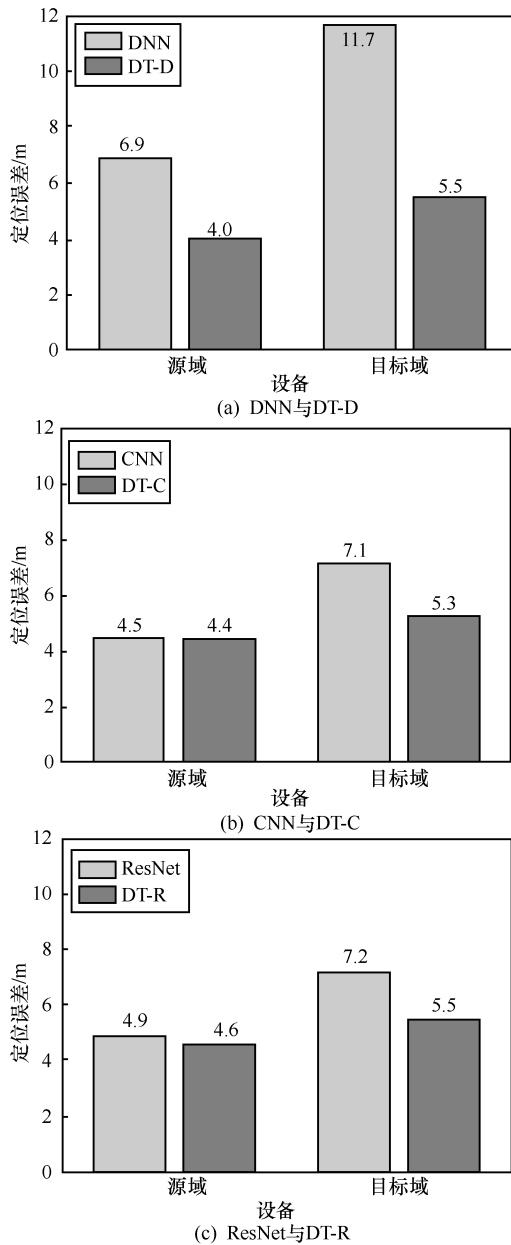


图8 商场数据集上 DeviceTransfer 设备无关性能的验证

由于在训练过程中仅使用源域数据的坐标标签, 且目标域设备与源域设备存在设备多样性, 因此在直接定位模型中, 在目标域设备上的定位精度比在源域设备上的定位精度差得多。而使用 Device Transfer 进行设备无关处理后, 在目标域设备上的定位精度均有明显提升, 且用 Device Transfer 后各个目标设备与源域设备的定位误差差距缩小。这是因为在设备无关特征学习的过程中, 特征提取器通过与解码器协作学习各个设备特定特征, 通过与生成器、鉴别器协作学

习跨设备共同特征, 然后在与定位器协作的过程中选择适当的特征进行定位。这说明设备无关的特征提取是有效果的, 能够有效减小不同设备之间的差距, 选取有助于定位的特征, 提高目标设备定位精度。经过 DeviceTransfer 域适应后, 源域设备的误差也在一定程度上降低, 说明模型提取的设备无关特征能有效帮助定位。基于以上分析, 本文验证了 DeviceTransfer 提取的设备无关特征的有效性。

定位器结构的选择。在本组对比实验中, 本文选择不同的网络结构作为定位器来验证 DeviceTransfer 的性能。根据实验结果, DeviceTransfer 在不同定位器上均能实现对目标设备定位精度的提升。以图 7 中 Redmi note 1S 为例, 此设备在 DT-D、DT-C、DT-R 上的定位误差分别为 0.8 m、0.9 m、0.9 m, 定位结果十分相近。观察其他设备也可以发现, 使用 DeviceTransfer 时, 同一设备在不同定位器上定位结果相差不大, 说明 DeviceTransfer 的整体框架对于设备无关特征提取的能力是强大的。又由于 DNN 的结构比 CNN 和 ResNet 简单, 因此本文选择 DNN 作为 DeviceTransfer 中的定位器。

2) DeviceTransfer 的定位性能验证

在此部分对比实验中, 仍选取 vivo X5M 手机作为源域设备, 其余手机作为目标域设备, 统计所有设备在不同定位模型中的定位误差。

DeviceTransfer 模型与基于对抗的 DANN 模型、CNNLoc 和设备无关定位算法 KNN-cosine 的对比如图 9 所示。以教学楼为例, 与自身模型在源域上的定位精度相比, DeviceTransfer 在 3 个目标域设备上的定位精度分别降低了 0.1 m、0.5 m、0.9 m; DANN 在 3 个目标域上的定位精度分别降低了 0.3 m、0.7 m、0.9 m; CNNLoc 在 3 个目标域上的定位精度分别降低了 0.3 m、0.6 m、0.9 m; KNN-cosine 在 3 个目标域上的定位精度分别降低了 0.4 m、0.8 m、1.4 m。在商场中, DeviceTransfer 在目标设备上的定位精度比 KNN-cosine、CNN Loc 和 DANN 在目标设备上的定位精度分别高 3.3 m、2.4 m、1.3 m。

因此, 与 DANN、CNNLoc、KNN-cosine 相比, DeviceTransfer 在所有的目标域设备上都可以取得最高的定位精度。

整体而言, DeviceTransfer 在 4 种模型中得到的定位结果最好, 在各个目标域上都能取得最高的定位精度。这说明 DeviceTransfer 有良好的定位性能, 在对不同设备做特征学习时学习设备无关的特征

能有效帮助定位器进行更准确的位置预测, 得到优于其他模型的定位结果; DeviceTransfer 具有较好的设备无关处理能力, 能够在跨设备数据集上取得良好的定位结果。

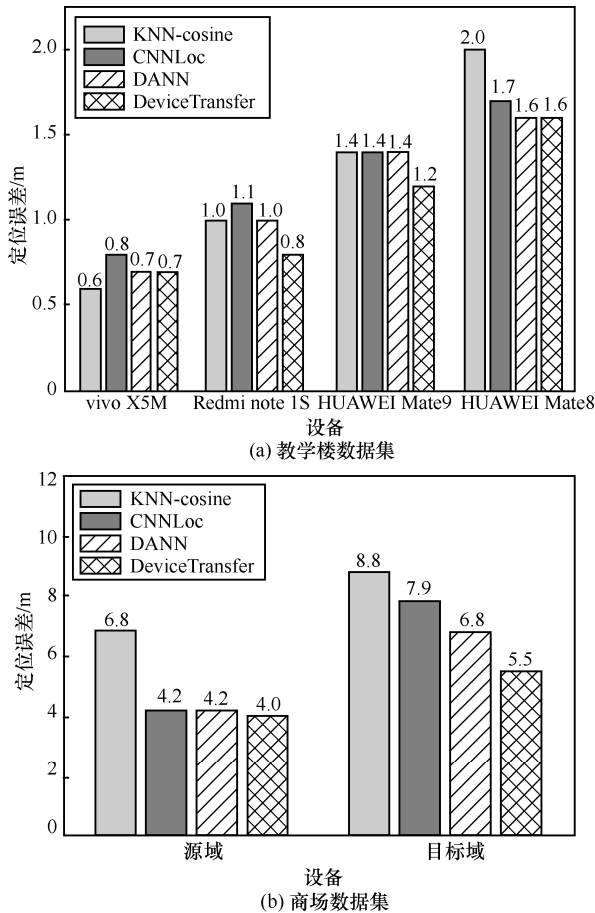


图 9 DeviceTransfer 模型与基于对抗的 DANN 模型、CNNLoc 和设备无关定位算法 KNN-cosine 的对比

5 结束语

基于 Wi-Fi 指纹的室内定位方法存在设备多样性问题, 严重影响了定位精度。基于此, 本文深入研究了域自适应方法, 以深度学习技术作为支撑, 设计了用于提取设备无关特征的模型 DeviceTransfer。该模型以设备类型为域, 提取不同设备 Wi-Fi 数据的设备无关特征, 再输入定位器进行位置预测, 能够把源域设备的位置信息转移到目标域, 有效克服设备多样性的影响, 获得较高的定位精度。该模型只需要源域设备在参考点上采集 Wi-Fi 数据 (即对 Wi-Fi 数据打上位置标签), 而目标域设备采集的 Wi-Fi 数据不需要位置信息, 因而极大降低了 Wi-Fi 数据的采集成本。

在模型设计过程中, 假设备类型的数量是有限的, 然而真实的应用场景中设备类型的数量往往是庞大的。因此, 为了提升模型的可扩展性, 后续的研究可以从增量学习的角度入手, 研究如何让模型在面对不断加入的新型设备时, 仍能提取设备无关的数据表示方法, 且不影响已有设备型号的定位效果。

参考文献:

- [1] HE S N, CHAN S H G. Wi-Fi fingerprint-based indoor positioning: recent advances and comparisons[J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 466-490.
- [2] 王慧强, 高凯旋, 吕宏武. 高精度室内定位研究评述及未来演进展望[J]. 通信学报, 2021, 42(7): 198-210.
WANG H Q, GAO K X, LYU H W. Survey of high-precision localization and the prospect of future evolution[J]. Journal on Communications, 2021, 42(7): 198-210.
- [3] LUI G, GALLAGHER T, LI B H, et al. Differences in RSSI readings made by different Wi-Fi chipsets: a limitation of WLAN localization[C]//Proceedings of 2011 International Conference on Localization and GNSS (ICL-GNSS). Piscataway: IEEE Press, 2011: 53-57.
- [4] ZHENG V W, PAN S J, YANG Q, et al. Transferring multi-device localization models using latent multi-task learning[C]//Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence. California: AAAI Press, 2008: 1427-1432.
- [5] TSUI A W, CHUANG Y H, CHU H H. Unsupervised learning for solving RSS hardware variance problem in Wi-Fi localization[J]. Mobile Networks and Applications, 2009, 14(5): 677-691.
- [6] HUANG C C, MANH H N, WANG Y S. An self-adaptive wireless indoor localization system for device diversity[C]//Proceedings of 2016 IEEE International Conference on Consumer Electronics. Piscataway: IEEE Press, 2016: 1-2.
- [7] ZHANG L Y, MA L, XU Y B, et al. Linear regression algorithm against device diversity for indoor WLAN localization system[C]//Proceedings of 2017 IEEE Global Communications Conference. Piscataway: IEEE Press, 2017: 1-6.
- [8] ZHANG L Y, MENG X L, FANG C. Linear regression algorithm against device diversity for the WLAN indoor localization system[J]. Wireless Communications and Mobile Computing, 2021, 2021: 5530396.
- [9] ZHAO H, DES-COMBES R T, ZHANG K, et al. On learning invariant representations for domain adaptation[C]//Proceedings of the 36th International Conference on Machine Learning. [S.l.:s.n.], 2019: 7523-7532.
- [10] YANG S, DESSAI P, VERMA M, et al. FreeLoc: calibration-free crowdsourced indoor localization[C]//Proceedings of IEEE INFOCOM. Piscataway: IEEE Press, 2013: 2481-2489.
- [11] CAI C W, DENG L, LI S F. CSI-based device-free indoor localization using convolutional neural networks[C]//Proceedings of 2018 IEEE 4th International Conference on Computer and Communications. Piscataway: IEEE Press, 2018: 753-757.
- [12] ASHRAF I, KANG M Y, HUR S, et al. MINLOC: magnetic field patterns-based indoor localization using convolutional neural networks[J]. IEEE Access, 2020, 8: 66213-66227.
- [13] WILSON G, COOK D J. A survey of unsupervised deep domain adaptation[J]. ACM Transactions on Intelligent Systems and Technology, 2020, 11(5): 1-46.

- [14] SUN B, FENG J, SAENKO K. Return of frustratingly easy domain adaptation[C]//Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. California: AAAI Press, 2016: 2058-2065.
- [15] KANG G L, JIANG L, YANG Y, et al. Contrastive adaptation network for unsupervised domain adaptation[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 4888-4897.
- [16] GHIFARY M, KLEIJN W B, ZHANG M, et al. Deep reconstruction-classification networks for unsupervised domain adaptation[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 597-613.
- [17] BOUSMALIS K, TRIGEORGIS G, SILBERMAN N, et al. Domain-separation networks[J]. Advances in Neural Information Processing Systems, 2016, 29: 343-351.
- [18] MATHUR A, ISOPOUSSU A, KAWSAR F, et al. FlexAdapt: flexible cycle-consistent adversarial domain adaptation[C]//Proceedings of 2019 18th IEEE International Conference on Machine Learning and Applications. Piscataway: IEEE Press, 2019: 896-901.
- [19] YI Z L, ZHANG H, TAN P, et al. DualGAN: unsupervised dual learning for image-to-image translation[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2868-2876.
- [20] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//Proceedings of the 32nd International Conference on Machine Learning. Cambridge: JMLR, 2015: 1180-1189.
- [21] ZHANG W C, OUYANG W L, LI W, et al. Collaborative and adversarial network for unsupervised domain adaptation[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3801-3809.
- [22] CHEN C H, MIAO Y S, LU C X, et al. Motion transformer: transferring neural inertial tracking between domains[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: ACM Press, 2019, 33: 8009-8016.
- [23] SOHN K, LEE H, YAN X. Learning structured output representation using deep conditional generative models[J]. Advances in neural information processing systems, 2015, 28: 3483-3491.
- [24] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. arXiv Preprint, arXiv:1411.1784, 2014.
- [25] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.
- [26] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2242-2251.
- [27] KINGMA D P, BA J. Adam: a method for stochastic optimization[J]. arXiv Preprint, arXiv: 1412.6980, 2014.
- [28] PASZKE A, GROSS S, CHINTALA S, et al. Automatic differentiation in pytorch[C]//Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017). New York: Curran Associates Inc, 2017: 1-4.
- [29] SONG X D, FAN X C, HE X J, et al. CNNLoc: deep-learning based indoor localization with WiFi fingerprinting[C]//Proceedings of 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCCom/IOP/SCI). Piscataway: IEEE Press, 2019: 589-595.
- [30] HAN S, ZHAO C, MENG W X, et al. Cosine similarity based fingerprinting algorithm in WLAN indoor positioning against device diversity[C]//Proceedings of 2015 IEEE International Conference on Communications. Piscataway: IEEE Press, 2015: 2710-2714.
- [31] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Cambridge: JMLR, 2015: 448-456.
- [32] XU B, WANG N, CHEN T, et al. Empirical evaluation of rectified activations in convolutional network[J]. arXiv Preprint, arXiv: 1505.00853, 2015.
- [33] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2010: 2528-2535.
- [34] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral normalization for generative adversarial networks[J]. arXiv Preprint, arXiv: 1802.05957, 2018.
- [35] HE K M, ZHANG X Y, REN S Q, et al. Identity mappings in deep residual networks[C]//Proceedings of the Computer Vision - ECCV 2016. Berlin: Springer, 2016: 630-645.

[作者简介]



赵增华（1974—），女，河南南乐人，天津大学副教授、硕士生导师，主要研究方向为移动和普适计算、室内定位、水下无线网络、软件定义无线网络、计算机网络协议和系统等。



童跃凡（1999—），女，福建莆田人，天津大学硕士生，主要研究方向为室内定位等。

崔佳洋（1998—），男，河北唐山人，天津大学硕士生，主要研究方向为室内定位等。