

基于扩散模型的室内定位射频指纹数据增强方法

艾浩军^{1,2}, 曾维珂^{1,2}, 陶荆杰^{1,2}, 徐锦盈^{1,2}, 常含笑^{1,2}

(1. 武汉大学国家网络安全学院, 湖北 武汉 430072; 2. 武汉大学空天信息安全与可信计算教育部重点实验室, 湖北 武汉 430072)

摘要: 射频指纹室内定位方法通过在离线阶段采集足量信号指纹建立密集指纹库保证定位精度。为降低指纹采集成本, 提出一种基于扩散模型的射频指纹数据增强方法 (FPDiffusion)。首先建立指纹序列的时序图表示, 通过高斯加噪方法实现扩散模型的前向过程, 反向过程采用 U 型自编码器网络, 根据射频指纹特性设计了网络的损失函数, 最后给出了基于稀疏指纹生成稠密指纹的计算过程。实验结果表明, 在仅有少量有标签指纹的情况下, FPDiffusion 方法在 K-近邻 (KNN) 和卷积神经网络 (CNN) 算法上的定位误差降低率分别达到 76% 和 28%, 在 KNN 上的定位精度相比高斯过程回归 (GPR) 和 GPR-GAN 方法有显著提升。

关键词: 扩散模型; 数据增强; 射频指纹; 室内定位

中图分类号: TP391

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2023200

Radio frequency fingerprint data augmentation for indoor localization based on diffusion model

AI Haojun^{1,2}, ZENG Weike^{1,2}, TAO Jingjie^{1,2}, XU Jinying^{1,2}, CHANG Hanxiao^{1,2}

1. School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China

2. Key Laboratory of Aerospace Information Security and Trusted Computing Ministry of Education, Wuhan University, Wuhan 430072, China

Abstract: The radio frequency fingerprint indoor localization method ensures the accuracy by collecting a sufficient amount of fingerprints in the offline state to build a dense fingerprint database. A data augmentation method called FPDiffusion was proposed based on diffusion model to reduce the cost of fingerprint acquisition. Firstly, a temporal graph representation of the fingerprint sequence was constructed, the forward process of the diffusion model was accomplished by adding Gaussian noise, and a U-Net was utilized for the reverse process. The loss function of the network was designed according to the characteristics of radio frequency fingerprints. Finally, the computational process for generating dense fingerprints based on sparse fingerprints was presented. Experimental results demonstrate that FPDiffusion achieves 76% and 28% localization error reduction on K-nearest neighbor (KNN) and convolutional neural network (CNN) respectively, and significantly improves localization accuracy on KNN compared to Gaussian process regression (GPR) and GPR-GAN when only a small amount of labeled fingerprints is available.

Keywords: diffusion model, data augmentation, radio frequency fingerprint, indoor localization

0 引言

室内定位技术可以在卫星导航定位系统失效的封闭室内场所提供导航位置服务。由于 Wi-Fi 和蓝牙设备的普及, 利用泛在射频信号的时空分布特性建立接

收信号强度指示 (RSSI, received signal strength indicator) 序列与位置的关系, 也就是指纹定位, 将机器学习方法用于解决位置预测问题^[1-3]。国内外学者提出了多种机器学习算法, 从经典的 K-近邻 (KNN) 到深度学习模型的引入, 这些方法取得了比较理想的结果。

收稿日期: 2023-06-29; 修回日期: 2023-10-10

基金项目: 国家自然科学基金资助项目 (No.61971316); 国家重点研发计划基金资助项目 (No.2016YFB0502204)

Foundation Items: The National Natural Science Foundation of China (No.61971316), The National Key Research and Development Program of China (No.2016YFB0502204)

在射频指纹室内定位领域,如何有效提取指纹数据特征对提升定位性能至关重要。传统的基于机器学习的定位方法对 RSSI 的波动性具有有限的表达性能,而深度学习的方法通过多层网络的叠加增强了对大型复杂数据的建模能力^[4]。因此,越来越多的学者将经典神经网络模型应用到指纹定位中。近年来,Transformer 结构^[5]的网络由于其优异的性能表现,逐渐成为各领域的主流模型结构。DRVAT (distributed representation vector and self-attention) 方法^[4]首次将 Transformer 结构和自注意力机制引入指纹定位中,并取得了可观的定位效果。

然而,同其他机器学习方法一样,射频指纹定位方法需要收集大量有标签指纹数据用于定位模型训练,以获得满意的定位性能,避免过拟合现象。而通过传统的人工数据采集方法获取有标签指纹需要大量的人力和成本。在机器学习领域中,数据增强作为一种低成本的有标签数据获取方法,被广泛应用以弥补训练数据的不足。同样地,数据增强也可以应用到基于机器学习的指纹定位中。

指纹数据增强的方法有 3 种。1) 基于传播模型的方法。由于射频信号传播的随机性,插值的方法并不能有效提升效果。为了生成未采集区域的指纹数据,Talvitie 等^[6]研究了一些常用的插值和外推技术并用于缺失位置的指纹生成,验证了适当的插值和外推方法可以提高定位精度。2) 随机加入方法。通过变换指纹中采纳的接入点 (AP, access point) 数目或在原指纹中添加随机噪声的方式获取变换的指纹^[7]。3) 基于生成模型的方法。深度学习尤其是生成对抗网络 (GAN, generative adversarial network)^[8]提出后,也有通过 GAN 基于原指纹数据库训练生成模型,再通过生成模型生成更多指纹的方法^[9]。Lan 等^[10]将指纹数据根据实际空间关系处理成图像形式并采用图像处理领域的超分辨率技术生成更细粒度的指纹数据; Sugasaki 等^[11]从类间学习 (BC, between-class learning) 方法推广出位置间 (BL, between-location) 数据增强方法,通过神经网络预测信号分布,再生成未采集位置的指纹。这些方法都在一定程度上降低了数据采集的工作量,维持了较好的精度。

基于深度学习的模型如 GAN^[8]、变分自编码器 (VAE, variational autoencoder)^[12]等已经在图像生成领域展现出优异的性能。对去噪扩散概率模型 (DDPM, denoising diffusion probabilistic model)^[13-14]的研究与应用取得了突破性的进展,在图像生成领

域的效果已经超过 GAN^[15],相关研究很好地证明了它的可扩展性和生成能力^[16-18]。由于信号在空间中的分布可以看作一幅二维的信号强度图像,图像生成的方法也可以运用到指纹生成领域。目前已有学者将 GAN、VAE 等方法应用于指纹生成场景中^[19],同样地,利用扩散模型优秀的数据生成能力,将其应用于射频指纹生成以降低数据采集成本具有不错的前景。

为了降低数据采集量,降低稀疏数据对定位精度的影响,本文提出了一种基于扩散模型的指纹数据增强方法 FPDiffusion (fingerprint diffusion)。首先,根据射频指纹信号的特点,提出了有利于扩散模型计算的射频指纹时序图表示方法;然后,详细地描述了射频指纹扩散模型,包括前向过程噪声的加入方法、以 U 型自编码器网络为基础的反向过程,设计了适应指纹定位的损失函数;最后,给出了利用扩散模型从已有指纹数据生成新数据的方法。实验对比了 GPR (Gaussian process regression)^[20]、GPR-GAN^[21]和 FPDiffusion 这 3 种增强方法的性能,结果表明 FPDiffusion 对 KNN 指纹定位方法的精度提升显著,对 CNN 定位方法的精度提升优于 GPR 且与 GPR-GAN 相当。

1 相关工作

1.1 指纹数据增强

数据增强方法可以分为有监督的方法和无监督的方法。有监督的数据增强方法使用预先设计好的规则在现有数据基础上进行变换,生成更多的数据,包括直接对单样本变换的方法和多样本合成的方法。单样本变换通常不会改变数据类别,也不需要修改标签。对于射频指纹生成,Sinha 等^[7]提出了 2 种单样本变换策略:一是对指纹中每个有效值分别减去固定值,使每个指纹样本可以生成与信号源数目相同的新样本;二是计算每个参考点 (RP, reference point) 上采集的各个信号源的信号强度均值,在生成新指纹时从初始指纹中每个信号强度值到其对应均值的范围中选取随机数作为新样本。

多样本数据增强方法使用多个样本来生成新的样本。例如,SMOTE (synthetic minority over-sampling technique) 方法^[22]通过插值的方式合成新样本来解决样本不平衡问题;SamplePairing 方法^[23]通过数据变换叠加方式生成新样本;Mixup 方法^[24]是基于邻域风险最小化原则的数据增强方法,通过 2 个随机

样本的加权叠加插值得到新数据及其标签。对于指纹生成, Sugasaki 等^[11]从 BC learning 方法推广出 BL 数据增强方法, 首先稀疏采集真实指纹, 统计各个信号源在同一位置的方差 (作为常数), 然后将坐标作为输入, 信号强度均值作为输出训练神经网络预测未采集位置的 RSSI 均值, 将对应位置的信号分布看作高斯分布, 最后对该分布采样得到未采集位置的指纹。

无监督的数据增强方法通过训练神经网络模型学习数据特征, 再进行数据生成。GPR-GAN 方法^[21]通过结合 GPR 模型和 GAN 模型进行指纹生成: 首先利用已采集的真实指纹训练 GPR 模型, 用于生成粗糙指纹作为 GAN 的输入, 训练对应的生成器和鉴别器; 最后通过生成器将 GPR 生成的目标位置的粗糙指纹变为接近真实数据的指纹。

1.2 扩散模型

扩散模型^[13-14]是一种用于图像去噪任务的机器学习模型, 旨在去除图像中的噪声, 同时保留重要特征。在扩散模型中, 模型通过应用一系列扩散步骤来逐渐改进噪声图像。前向过程中每个扩散步骤按固定策略向图像中添加噪声。反向过程使用神经网络进行去噪处理, 重复多次使每个步骤中的噪声水平逐渐降低, 直到得到最终的去噪图像, 模型使用原图像和噪声图像组成的图像对进行训练, 训练过程中学习噪声的分布和去噪过程的参数, 捕捉数据中的统计模式,

使其能够有效对任意图像去噪。

扩散模型采用概率建模, 模型为给定的噪声输入生成一系列可能的去噪图像分布, 允许进行不确定性估计和基于采样的图像生成, 对于图像修复、超分辨率等领域都有很高的应用价值。扩散模型作为一种通用的概率扩散模型, 也可以应用到其他领域的任务中, 如在自然语言处理、音频处理中用于建模数据的分布和生成可能的样本。本文尝试将扩散模型用于射频指纹的生成, 通过少量样本扩充指纹数据库提升定位精度。

2 射频指纹数据增强方法

2.1 方法概述

本文以扩散模型理论为基础, 提出了一种新的射频信号指纹数据增强方法 FPDiffusion, 通过稀疏采集的有标签指纹生成更多目标位置的指纹。方法整体结构如图 1 所示, 主要包括以下 3 个部分。

1) 射频指纹数据预处理。提出了射频时序图表示方法, 将采集的众源指纹数据和有标签指纹数据以二维方式组织, 进行归一化。

2) 扩散模型训练。使用以二维方式组织的众源数据训练 U 型自编码器网络。通过随机选择扩散步骤数 t 的方式, 生成原始射频时序图在步骤数 t 时对应的高斯噪声图像和噪声对应的均值及方差, 并将生成的噪声图像作为网络输入, 生成的均值和方

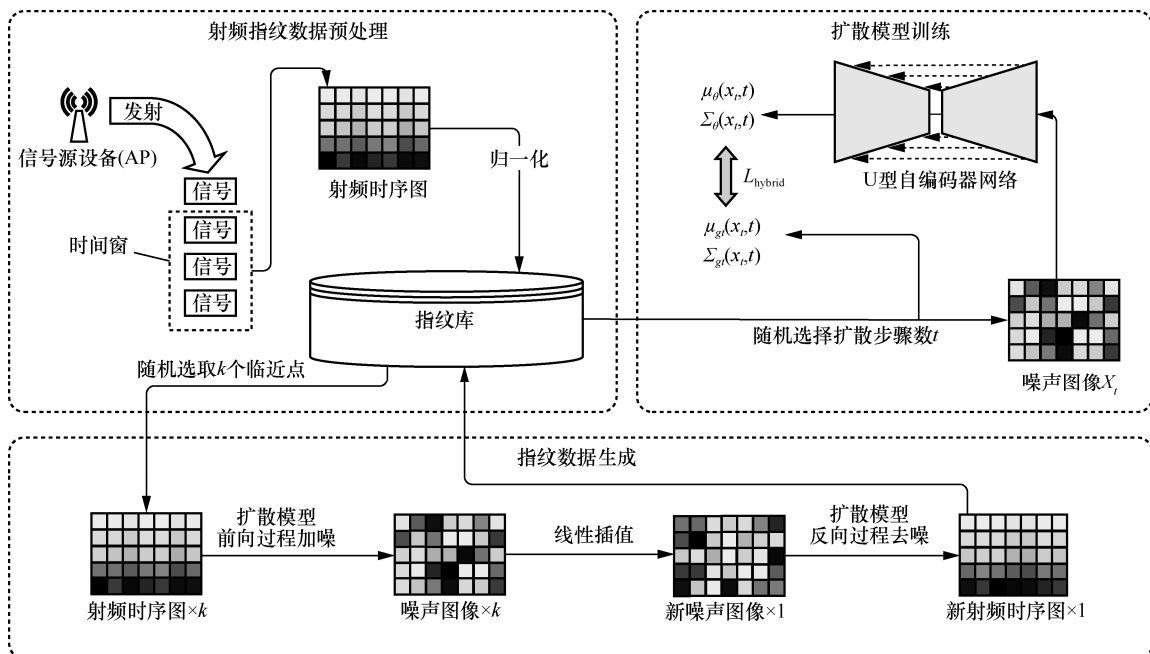


图 1 射频指纹数据增强方法整体结构

差作为真值，最终的训练目标是使网络预测的均值和方差尽可能地逼近真值。

3) 指纹数据生成。从随机选取的临近点中各抽取一张射频时序图，通过扩散模型的前向过程逐步加噪，将得到的噪声图像编码到隐空间中并进行线性插值，生成目标位置指纹数据对应的噪声图像。通过反向过程训练好的 U 型自编码器网络对噪声图像逐步去噪，获得目标位置指纹数据对应的射频时序图，加入指纹库中。

2.2 射频指纹数据结构构建

2.2.1 射频时序图

移动终端捕捉到的蓝牙射频信号是一维时间序列，一个时间序列内通常包含来自同一基站的多个样本，为此提出一种有利于以图像方式为蓝牙射频信号建模的数据组织方法，即“射频时序图”，如图 2 所示。一张射频时序图包含一个时间窗内接收到的所有信号源的 RSSI 信息，其中每列对应一个信号源设备，列中每个数据为按时间顺序接收到的信号强度信息。纵向维度设置为可容纳有效信息的长度，对于接收到的蓝牙广播包数目较少的设备，根据 RSSI 的实际取值范围选取一个下边界值作为默认填充值(图 2 中 PAD 所示)将多余的空位进行填充，表示无信息数据位。通过该方法组织的射频时序图作为蓝牙指纹，保留了蓝牙广播数据不同信道的差异性特征，将其组织成二维形式并引入时序关系，使数据更方便地适配扩散模型并具备更加丰富的信息维度。

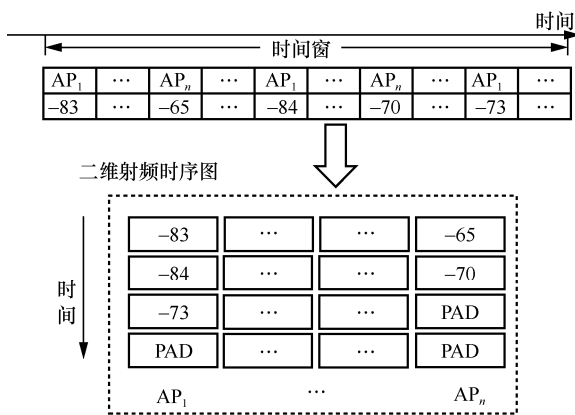


图 2 射频时序图组织方法

2.2.2 指纹数据归一化

为使信号强度数据符合扩散模型反向过程中从标准高斯分布降噪采样的数据适宜范围，将其进行归一化。根据室内环境中真实数据的信号强度范

围进行归一化范围的选取，将其线性缩放到[-1,1]。记选取的区间左边界为 R_{min} ，右边界为 R_{max} ，将区间范围内的信号强度 R 归一化为 $S(R)$ ，计算方法如下

$$S(R) = -1 + \frac{R - R_{min}}{R_{max} - R_{min}} \times 2 \quad (1)$$

在射频时序图中，除有效信号强度外，还有用于填充的 padding 值，需选取一个实际数值与之对应。考虑到 padding 值表示“无信号数据”，其在数值上应与极弱信号强度相对应，令 padding 对应归一化后的边界值-1。在选取信号强度范围时，将有效区间左端进行适当左移，使正常弱信号强度值不会到达端点，与 padding 值进行区分。对于部分超出选取区间的离群值，使用边界取值作为缩放结果。射频时序图中信号强度的缩放示例如图 3 所示。

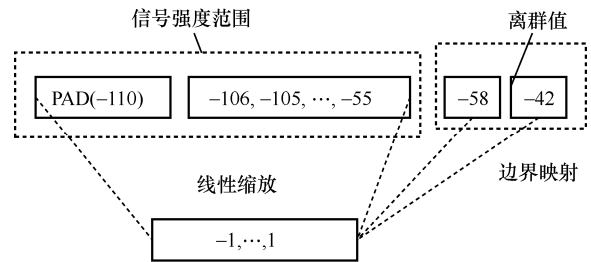


图 3 信号强度缩放示例

2.3 射频指纹扩散模型

2.3.1 基本过程概述

FPDiffusion 模型的基本过程如图 4 所示，包括前向加噪过程和反向去噪过程。前向加噪过程采用余弦式的噪声策略向原始输入的射频时序图 x_0 中逐步添加噪声，至 T 步时射频时序图变为近似标准高斯噪声图像。反向去噪过程从标准高斯噪声图像开始逐步去噪，去噪方法是训练 U 型自编码器网络预测前一步的均值和方差，经过 T 步去噪最终得到清晰的射频时序图 x_0 。

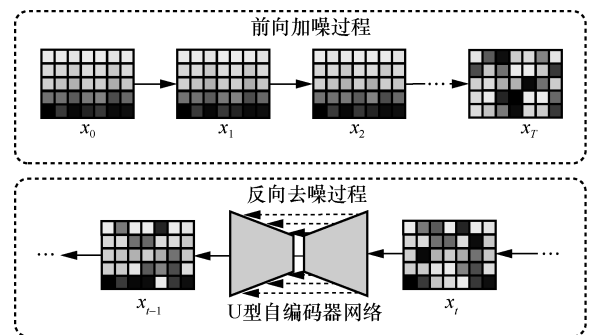


图 4 FPDiffusion 模型基本过程

2.3.2 前向加噪过程

FPDiffusion 模型的前向加噪过程采用余弦式的噪声策略^[25]，其计算方式如下

$$\bar{\alpha}_t = \frac{f(t)}{f(0)}, f(t) = \cos\left(\frac{\frac{t}{T} + s}{1 + s} \frac{\pi}{2}\right)^2 \quad (2)$$

其中， t 表示当前扩散步骤数，前向加噪过程总扩散步骤数为 T 步。 s 是一个小的偏移量，目的是防止前向加噪过程初始轮次的方差过小。该策略以 $\bar{\alpha}_t$ 为中心，计算方差参数 β_t 的方式为

$$\beta_t = 1 - \frac{\alpha_t}{\alpha_{t-1}} \quad (3)$$

将初始输入的射频时序图表示为 x_0 ，设 x_1, \dots, x_T 为和 x_0 相同维度的隐变量，表示扩散过程中第一步~第 T 步产生的噪声图像。按照拟定的方差参数 β_1, \dots, β_T 逐步向输入的射频时序图添加高斯噪声的过程可表示为

$$q(x_{1:T} | x_0) := \prod_{t=1}^T q(x_t | x_{t-1})$$

$$q(x_t | x_{t-1}) := N(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \quad (4)$$

由此可知，前向加噪过程的每一步都是对上一步得到的噪声图像 x_{t-1} 按照拟定的方差参数 β_t 添加高斯噪声，得到的新噪声图像 x_t 都是由 x_{t-1} 和 β_t 决定的高斯分布。

2.3.3 反向去噪过程

1) 网络结构

U 型自编码器网络被 Pix2Pix GAN 方法^[26]首次运用到深层次生成模型中，证明了其强大的生成能力。FPDiffusion 模型反向去噪过程的每一步均采用 U 型自编码器网络预测前一步的均值和方差。整体网络结构如图 5 所示，为适配射频时序图的数据特点，使用更少的下采样深度和通道数。

① 编码器。逐层提取输入噪声图像的特征，将原始的高维数据编码到低维隐空间中。二维卷积提升输入噪声图像的通道数，使模型从更多角度观测数据。由 Res 块和自注意力块构成的特征提取块对当前维度下的特征进行计算提取。下采样层将噪声图的宽度和高度均缩小一半，通道数提升。最后一个特征提取块的输出直接连接到中间特征层以防止过多下采样使噪声图像数据维度过小。

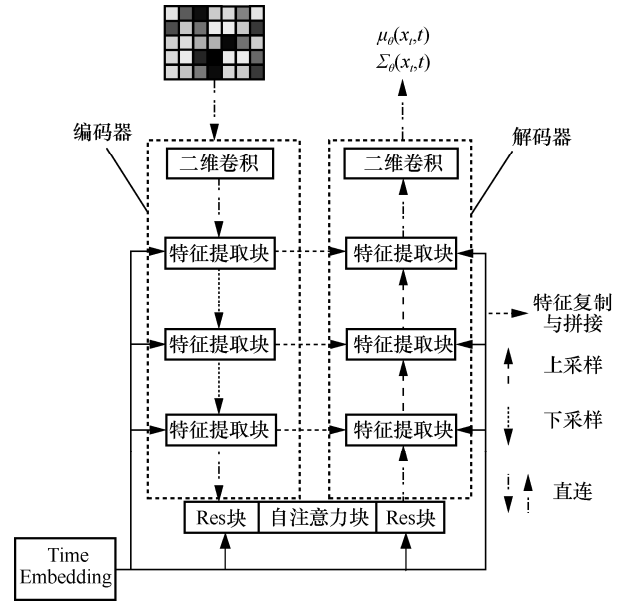


图 5 U 型自编码器网络结构

② 中间特征层。从经过编码器降维后的低维度特征中提取高级特征，作为解码器的输入。

③ 解码器。通过两次上采样将数据维度恢复到噪声图初始大小。上采样层将噪声图像已编码特征的宽度和高度增加一倍，并减少通道数。输出层通过二维卷积调整通道数使数据结构和输出目标一致，即预测的均值和方差。

④ 跳跃连接。由于噪声图像的长和宽具有不规则性，编码器下采样过程中会出现奇数维度的输入，与经下采样降维取整后再上采样得到的偶数维度不对应，故进行跳跃连接时要根据实际维度进行裁剪后再拼接。

⑤ Time Embedding 层。由于网络输入的噪声图像与输出的均值和方差需要随着扩散步骤数 t 变化，为使同一神经网络能够作用于反向去噪过程的不同步骤，根据当前步骤数 t 从不同尺度处理数据，将扩散步骤数 t 引入作为网络输入。

该层由使用 Sigmoid 激活函数的线性单元构成，将步骤数 t 转换为和射频时序图数据维度对应的高维向量。网络中每个特征提取块的输入都与转换过的 Time Embedding 叠加后再进行特征提取，使特征处理在步骤数 t 下进行。

2) 训练目标与损失函数

FPDiffusion 模型的总损失函数为

$$L_{\text{hybrid}} = L_{\text{simple}} + \lambda L_{\text{VLB}} \quad (5)$$

其中，权重 λ 是预先设置的超参数； L_{simple} 是简化的损

失函数，使模型更容易训练； L_{VLB} 是包含方差项的完整损失函数，由于对数似然的微小提升可以对样本生成质量和特征表示的学习产生很大的影响^[27]，通过该损失项将模型预测的方差 $\sum_{\theta} (x_i, t)$ 加入训练过程，并

使用 Nichol 等^[25]提出的对数域插值方式表示方差，即

$$\sum_{\theta} (x_i, t) = \exp(v \log \beta_i + (1-v) \log \tilde{\beta}_i) \quad (6)$$

损失项 L_{simple} 和 L_{VLB} 的详细计算过程见附录 1 和附录 2。

2.4 指纹生成方法

考虑在大面积室内定位场景内均匀划分网格、间隔采集有标签指纹的采集方案，对于其中任意非边缘点，取 k 个近邻点，采用 K 近邻插值方法得到加权系数。下面以 $k=4$ 的情况，详细描述生成过程，如图 6 所示。其中， \oplus 表示线性插值操作。

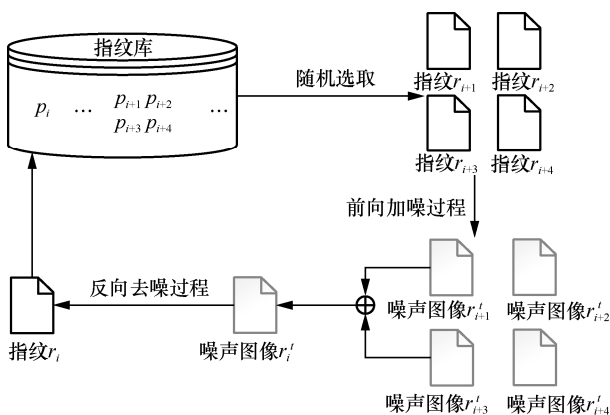


图 6 指纹生成流程

为方便表述，进行如下符号定义。记非边缘待生成参考点为 p_i ，射频时序图为 r_i ，噪声图为 r'_i 。

1) 对于任意 p_i ，取其四周近邻的已采集指纹参考点，分别表示为 $p_{i+1}, p_{i+2}, p_{i+3}, p_{i+4}$ ，且 $p_i \neq p_{i+1} \neq p_{i+2} \neq p_{i+3} \neq p_{i+4}$ ，设点 p_i 与 $p_{i+1}, p_{i+2}, p_{i+3}, p_{i+4}$ 的距离分别为 d_1, d_2, d_3, d_4 ，其位置关系如图 7 所示。从稀疏采集的指纹库中各随机抽取一张 $p_{i+1}, p_{i+2}, p_{i+3}, p_{i+4}$ 对应的射频时序图 $r_{i+1}, r_{i+2}, r_{i+3}, r_{i+4}$ 。

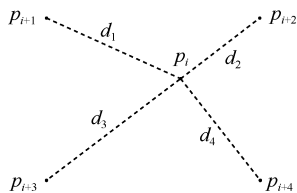


图 7 参考点位置关系

2) 使用扩散模型前向加噪过程对 $r_{i+1}, r_{i+2}, r_{i+3}, r_{i+4}$ 进行加噪编码得到噪声图像 $r'_{i+1}, r'_{i+2}, r'_{i+3}, r'_{i+4}$ 。

3) $r'_{i+1}, r'_{i+2}, r'_{i+3}, r'_{i+4}$ 按比例叠加得到一张待生成参考点 p_i 隐空间编码的噪声图像 r'_i ，即

$$r'_i = R \left(\frac{1}{d_1} r'_{i+1} + \frac{1}{d_2} r'_{i+2} + \frac{1}{d_3} r'_{i+3} + \frac{1}{d_4} r'_{i+4} \right) \quad (7)$$

为保证每一项的系数和为 1，定义 R 为比例修正系数，表示为

$$R = \frac{1}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \frac{1}{d_4}} = \frac{d_1 d_2 d_3 d_4}{d_2 d_3 d_4 + d_1 d_3 d_4 + d_1 d_2 d_4 + d_1 d_2 d_3} \quad (8)$$

4) 将按比例叠加得到的噪声图 r'_i 作为反向去噪过程 U 型自编码器网络的输入，对其进行解码去噪，得到待生成参考点 p_i 对应的射频时序图 r_i 作为生成的指纹，指纹标签即点 p_i 的坐标。

重复上述步骤直至生成的指纹数量足够，对室内空间中的其余参考点继续该过程，直至所有点都获得足量指纹。

2.5 指纹距离计算方法

数据生成领域目前缺乏合适的评估标准，往往只对数据增强后的定位效果进行验证，以说明方法的合理性。因此，本文提出一种射频指纹相似度对比方法，通过计算生成指纹与真实指纹之间的距离来评价生成指纹的优劣。

对于 AP 平均映射的指纹表示方式，指纹数据被按照信号来源组织成简单的一维向量，因此可以直接计算 2 个向量的欧氏距离或者余弦距离来表示 2 个指纹数据的相似度。因为在同样时间窗内，不同 AP 发出的信号被捕获的数量不一样，所以在射频时序图中，从第二行开始有些 AP 的强度数据使用了默认填充值 PAD，如图 8 所示。这样直接计算距离是不合理的。

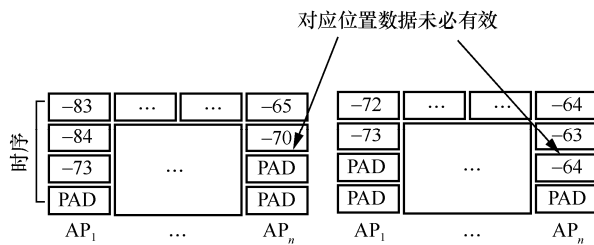


图 8 射频时序图坐标对应关系

为此本文提出一种计算射频时序图距离 D 的方法如下

$$D = \frac{\sum Dcol^i}{n} \quad (9)$$

$$Dcol^i = \min \left(\sqrt{\sum_j \frac{(R_{i,j}^1 - R_{i,j}^{2shift})^2}{\alpha^{\max(0, j-k)}}} \right) \quad (10)$$

其中， $Dcol^i$ 为 2 个射频时序图第 i 列的距离； n 为预设信号源的设备数，即射频时序图的列数； α 和 k 为预设的常数， k 为在计算距离时这一列纳入完全考虑（不进行缩放）的元素数量， α 为缩放参数，表示在超过 k 个元素之后的每个元素在计算距离后其计入总距离的权重按指数进行缩放； $R_{i,j}^1$ 表示第一个射频时序图第 i 列的第 j 个元素； $R_{i,j}^{2shift}$ 表示第二个射频时序图经过数据平移后的第 i 列的第 j 个元素。

数据平移方式是在一列有效数据中进行平移，将末尾的有效数据拼接到该列首部，并将其他数据后移，如图 9 所示。其目的是考虑到射频时序图的一列中的时序关系，在进行平移时保持部分数据的连续关系不变。由于射频时序图中的信号强度数据可能来自不同的蓝牙广播信道，将数据平移是在尝试将时序数据的起点信道对齐，以缩小因信道差异导致的计算距离。此处计算列距离时将变换过的多组计算结果进行对比，取最小的一组作为最终的列距离。

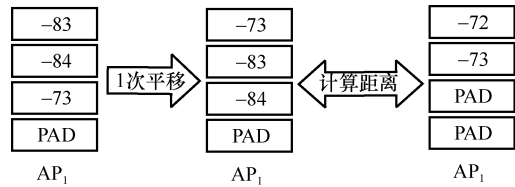


图 9 列数据平移示例

加入缩放参数的目的是缩小由设备在不同时间窗下接收的广播包数目差距带来的影响。对于前 k 个数据包通常都可以接收到，在计算时不进行缩放，而第 $k+1$ 及之后的数据包可能是无效数据，因此使用缩放参数将该部分对距离计算的影响减小，并且越靠后的数据缩放倍数呈指数级增加。

3 实验与分析

3.1 室内实验环境与数据采集

室内实验环境为横向 17.5 m，纵向 6.7 m 的矩形办公实验室，包含桌椅、电脑等物品并有日常人员走动，室内信号环境相对稳定且具备日常干扰波动，具有一般代表性，如图 10 所示。其中，星形图标表示 19 个部署的蓝牙基站，向外发射蓝牙广播信号，供接收端用于指纹定位；圆形图标表示按 1 m 间隔划分的参考点。采集数据的方式分为 2 种。

一是在参考点上采集数据。使用手机接收来自各个信号源的广播数据包并记录 MAC 地址、信号强度等重要信息，将参考点坐标作为相应数据的位置标签。使用该种方式采集的有标签稀疏数据集用

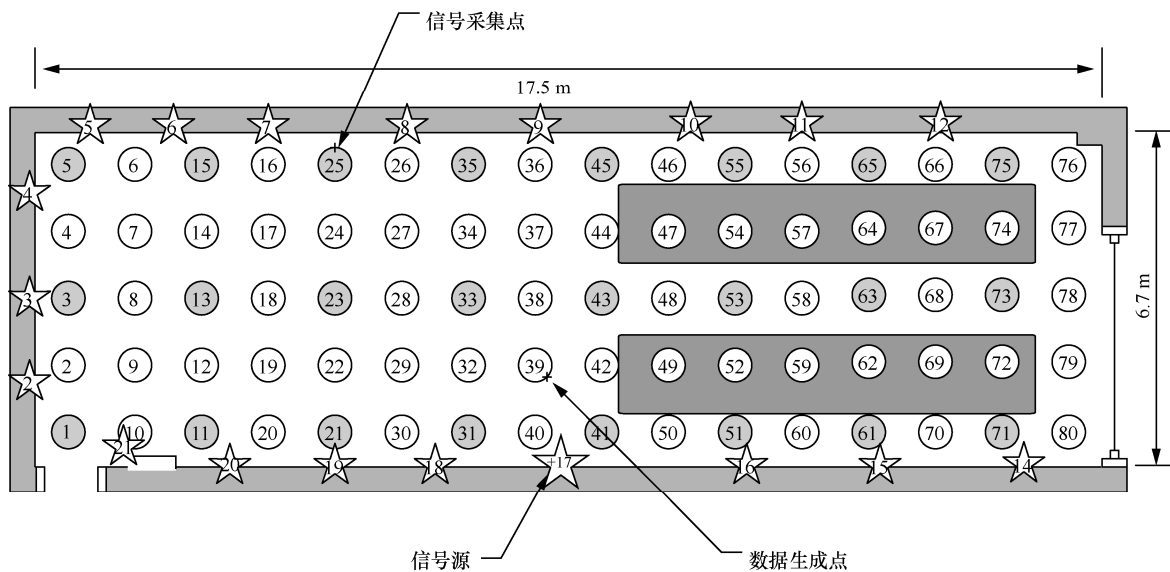


图 10 信号采集环境的室内地图与采集方案

于定位模型的训练和评估。训练集和测试集采集方式相同，即在所有参考点上各采集 3 min。由于随时间推移，室内环境改变会引起射频信号特征变化，设置较长的采集间隔时间以模拟真实的室内定位情况，即测试集的采集时间在训练集采集间隔一个半月之后。

二是手持接收设备在室内随机行走收集大量众源数据。众源数据不包含位置标签，因此不能直接用于训练定位模型，但可以 and 定点采集的有标签数据一同用作模型的预训练，丰富预训练数据库。

对于 FPDiffusion 方法，训练扩散模型时使用众源数据，指纹生成时使用有标签数据进行辅助，采用间隔测量、中间插值的方式，即选取图中灰色参考点的数据作为基准，辅助生成白色参考点的数据。

3.2 实验设置

1) 扩散模型参数设置

射频指纹扩散模型实验参数设置如表 1 所示。

表 1 射频指纹扩散模型实验参数设置

参数符号	参数取值	参数说明
β_t	0.999	前向过程中，为防止结尾的一些轮次方差过大设置的一个上限值
s	0.008	前向过程余弦式噪声策略中的小偏移量
λ	0.001	反向过程总损失函数中 L_{VLB} 的权重，防止方差对模型训练的影响过大，同时对 L_{VLB} 应用梯度停止策略
ep	47 000	反向过程网络训练轮次
lr	1×10^{-4}	学习率
bs	512	批量大小
T	1 000	扩散步骤数最大值

2) 定位模型设置

使用 KNN 和 CNN 作为定位模型来评估指纹数据增强方法的性能。KNN 中的 k 设置为 3，进行定位时，根据实时指纹在有标签指纹库中进行相似度匹配，取前 k 个最邻近的指纹对应坐标进行加权平均。CNN 包括 2 个卷积层和 3 个全连接层，利用有标签指纹进行训练，定位时将实时指纹输入训练好的模型预测位置坐标。输入的射频时序图尺寸为 10 像素×19 像素。

3) 扩散步骤数的选取

扩散步骤数 t 需要进行合适的选取以获得合

理的插值，在去噪充分的情况下防止插值指纹变化为不可控样本。使用指纹距离热力图显示生成指纹的距离与反向过程轮次的关系，选择 18 号参考点作为插值目标点，从 13 号和 23 号参考点的指纹数据中抽取样本，通过 2.4 节提出的指纹生成方法生成 18 号的指纹，再与其他实际指纹进行距离对比。取扩散步骤数 $t=100、125、250、375、500、625、750、875、1\ 000$ ，结果如图 11 所示。

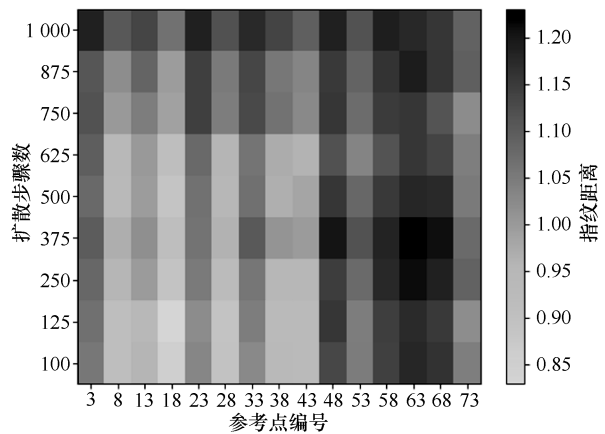


图 11 不同扩散步骤数生成的 18 号参考点指纹与真实指纹间距离

热力图的每一行中，18 号参考点对应的方块几乎都是最亮的位置，证明 FPDiffusion 方法生成的指纹与真实值较接近。扩散步骤数 $t=1\ 000$ 时，生成指纹与各个参考点的真实指纹距离都较远，可以认为是生成了不依赖于源数据的全新样本。扩散步骤数从 1 000 降低时，18 号参考点对应列的亮度呈升高趋势，取 125 时，生成指纹与 18 号参考点实际指纹的距离达到最小值，并且与其他参考点的真实指纹保持较大差异，此时生成质量最佳。

取扩散步骤数 $t=100、125、250、500、750、1\ 000$ ，生成所有未采集参考点的指纹，加入稀疏训练集训练 CNN 模型进行定位精度测试，平均指纹距离和平均定位误差 (ALE, average localization error) 随扩散步骤数的变化趋势如图 12 所示。

由图 12 可知，平均指纹距离和平均定位误差随扩散步骤数的变化趋势基本保持一致，证明了指纹距离作为指纹生成质量评价指标的有效性，能够反映数据增强后定位精度的变化情况。图 12 中指纹距离和定位误差在 125 步处均达到最小值，因此

综合以上分析，选取 $t=125$ 作为最佳扩散步骤数，并作为后续实验的超参数。

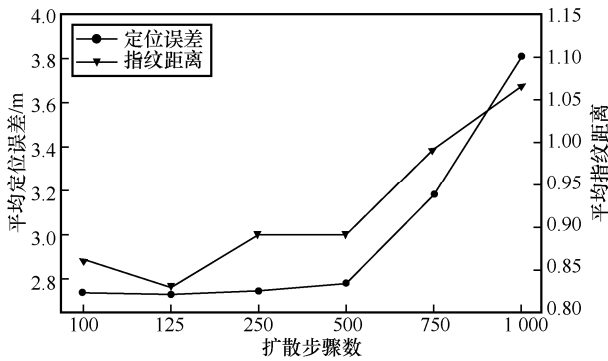


图 12 不同扩散步骤数进行数据增强的指纹距离和定位误差折线图

3.3 定位效果分析

选择定位误差累积分布函数 (CDF, cumulative distribution function) 表征各组定位结果。使用经 FPDiffusion 方法增强后得到的有标签密集指纹库训练室内定位模型，其定位精度作为数据增强效果的展示组，设置对照组：使用稀疏指纹（即图 10 中灰色参考点上采集的所有指纹）训练的室内定位模型、使用真实的完整密集指纹训练的室内定位模型。为排除众源数据训练扩散模型导致展示组获得更多数据信息的影响，所有室内定位模型都使用同样的众源数据进行预训练。

各对照组在训练集下训练的 KNN 和 CNN 定位结果分别如图 13 和图 14 所示。经过 FPDiffusion 方法数据增强后训练的室内定位模型精度显然高于只使用稀疏指纹训练的模型，证明了本文方法的有效性。虽然尚未达到真实密集采集数据的定位精度，但对稀疏数据场景而言已具有不错的提升。

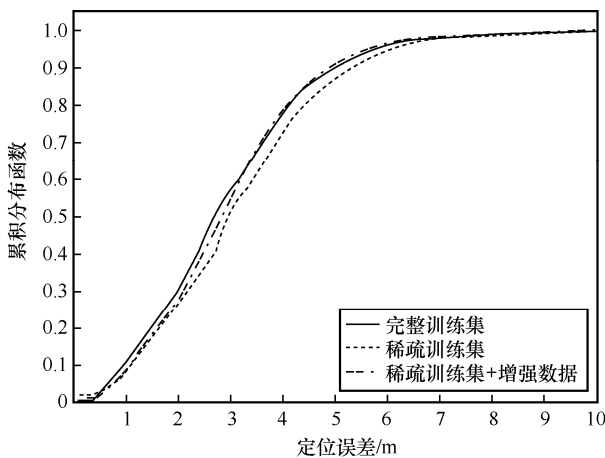


图 13 在不同训练集下训练的 KNN 定位结果对比

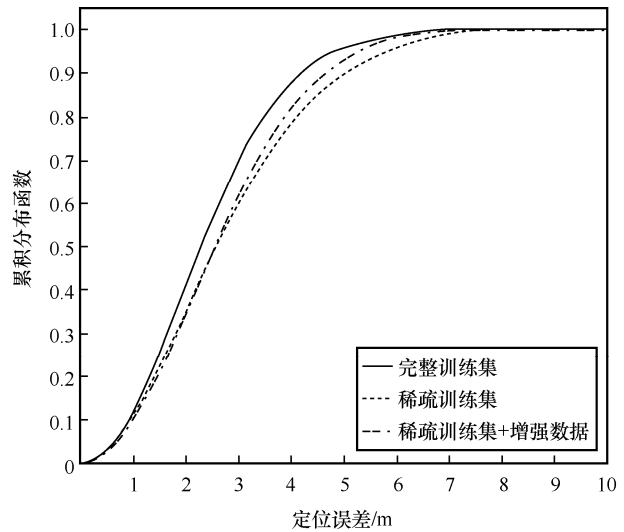


图 14 在不同训练集下训练的 CNN 定位结果对比

可以计算稀疏数据集和增强后密集数据集的定位误差差异，也就是误差降低率 Rd ，定量评价 FPDiffusion 方法对定位精度的提升程度。记使用稀疏数据集的平均定位误差为 E_l ，使用真实密集数据集的平均定位误差为 E_m ，使用 FPDiffusion 方法的平均定位误差为 E_s ，则误差降低率 Rd 计算方法如下

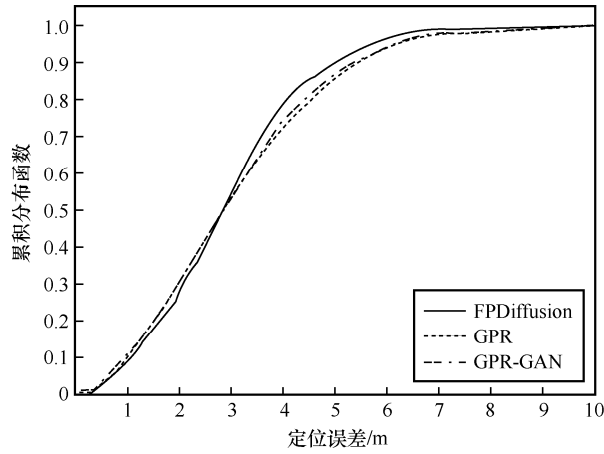
$$Rd = \frac{E_l - E_s}{E_l - E_m} \times 100\% \quad (11)$$

使用不同数据集的 ALE 以及根据式(11)计算 FPDiffusion 方法在各定位模型下的 Rd 如表 2 所示。

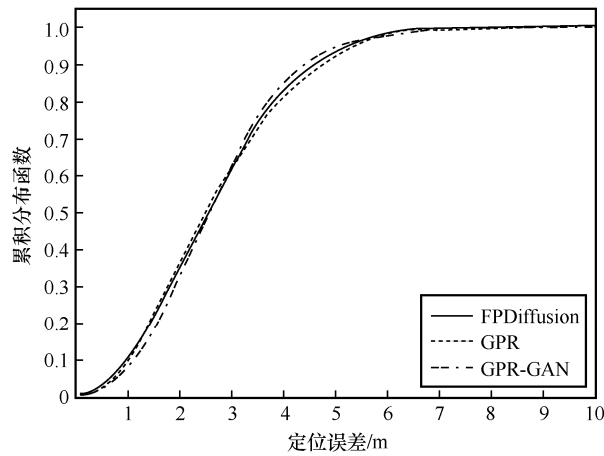
表 2 使用不同数据集的 ALE 和 FPDiffusion 的 Rd

定位模型	ALE/m			Rd
	E_l	E_m	E_s	
KNN	3.17	2.96	3.01	76%
CNN	2.83	2.47	2.73	28%

将本文提出的 FPDiffusion 方法与 GPR^[20]和 GPR-GAN^[21]指纹数据增强方法应用在前述稀疏数据集上进行对比，选择 KNN 和 CNN 定位模型验证其定位精度提升程度。3 种指纹增强方法的 CDF 对比如图 15 所示，ALE 和 Rd 对比如表 3 所示。实验结果表明，FPDiffusion 方法在 KNN 模型上的表现优势明显，在 CNN 模型上的表现优于 GPR 并与 GPR-GAN 相当。



(a) 3种指纹增强方法在KNN中的定位结果对比



(b) 3种指纹增强方法在CNN中的定位结果对比

图 15 在稀疏数据集上应用 3 种增强方法的定位结果对比

表 3 应用 3 种指纹增强方法的 ALE 和 Rd 对比

定位模型	GPR		GPR-GAN		FPDiffusion	
	ALE/m	Rd	ALE/m	Rd	ALE/m	Rd
KNN	3.14	14%	3.12	24%	3.01	76%
CNN	2.74	25%	2.73	28%	2.73	28%

4 结束语

针对射频指纹采集困难导致定位精度受限的问题，提出了一种基于扩散模型的射频指纹数据增强方法 FPDiffusion。实验结果表明，本文方法将扩散模型与 U 型自编码器网络相结合，能够有效学习多径效应对信号分布的潜在影响，对 KNN 模型的定位误差降低率达到 76%，对 CNN 模型的定位误差降低率达到 28%。对比 GPR 和 GPR-GAN 数据增强方法，在相同的标记样本数量下，显著地提升了 KNN 模型的定位精度。未来研究中，可将射频信号的室内传播分布规律用于引导扩散模型的生

成过程，使增强的指纹数据更符合信号的时空分布，进一步提升生成指纹的定位精度。

附录 1 L_{VLB} 损失项

FPDiffusion 模型的反向过程可看作前向过程的逆过程，模型训练过程中负责尽可能准确地预测前向过程，通过优化负对数似然的变分下界来进行训练

$$\mathbb{E}[-\log p_{\theta}(x_0)] \leq \mathbb{E}_q \left[-\log \frac{p_{\theta}(x_0)}{q(x_{1:T} | x_0)} \right] = \mathbb{E}_q \left[-\log p(x_T) - \sum_{t \geq 1} \log \frac{p_{\theta}(x_{t-1} | x_t)}{q(x_t | x_{t-1})} \right] =: L \quad (12)$$

因此 L_{VLB} 定义为

$$L_{\text{VLB}} = \mathbb{E}_{q(x_{0:T})} \left[\log \frac{q(x_{1:T} | x_0)}{p_{\theta}(x_{0:T})} \right] \quad (13)$$

其中，联合分布 $p_{\theta}(x_{0:T})$ 定义为从标准高斯分布 $p(x_T) = N(x_T; 0, I)$ 开始的马尔可夫链，即

$$p_{\theta}(x_{0:T}) := p(x_T) \prod_{t=1}^T p_{\theta}(x_{t-1} | x_t) \quad (14)$$

$$p_{\theta}(x_{t-1} | x_t) := N(x_{t-1}; \mu_{\theta}(x_t, t), \Sigma_{\theta}(x_t, t)) \quad (15)$$

根据式(4)、式(14)进一步变换 L_{VLB} 可得

$$L_{\text{VLB}} = \mathbb{E}_q \left[\log \frac{\prod_{t=1}^T q(x_t | x_{t-1})}{p_{\theta}(x_T) \prod_{t=1}^T p_{\theta}(x_{t-1} | x_t)} \right] = \mathbb{E}_q \left[\log \frac{q(x_T | x_0)}{p_{\theta}(x_T)} - \sum_{t=2}^T \log \frac{q(x_{t-1} | x_t, x_0)}{p_{\theta}(x_{t-1} | x_t)} - \log p_{\theta}(x_0 | x_1) \right] = \mathbb{E}_q \left[\underbrace{D_{\text{KL}}(q(x_T | x_0) \| p_{\theta}(x_T))}_{L_T} + \sum_{t=2}^T \underbrace{D_{\text{KL}}(q(x_{t-1} | x_t, x_0) \| p_{\theta}(x_{t-1} | x_t))}_{L_{t-1}} - \underbrace{\log p_{\theta}(x_0 | x_1)}_{L_0} \right] \quad (16)$$

由此可见，该损失项可以变为多项 KL 散度和熵的叠加。通过 KL 散度直接比较 $p_{\theta}(x_{t-1} | x_t)$ 与前向过程 $q(x_t | x_{t-1})$ 的后验

$$q(x_{t-1} | x_t, x_0) = q(x_t | x_{t-1}, x_0) \frac{q(x_{t-1} | x_0)}{q(x_t | x_0)} \propto \exp \left[-\frac{1}{2} \left(\frac{(x_t - \sqrt{\alpha_t} x_0)^2}{\beta_t} + \frac{(x_{t-1} - \sqrt{\alpha_{t-1}} x_0)^2}{1 - \alpha_{t-1}} - \frac{(x_t - \sqrt{\alpha_t} x_0)^2}{1 - \alpha_t} \right) \right] =$$

$$\exp \left(-\frac{1}{2} \left(\underbrace{\left(\frac{\alpha_t}{\beta_t} + \frac{1}{1-\alpha_{t-1}} \right) x_{t-1}^2}_{x_{t-1} \text{ 方差}} + \underbrace{\left(\frac{2\sqrt{\alpha_t}}{\beta_t} + \frac{2\sqrt{\alpha_{t-1}}}{1-\alpha_{t-1}} x_0 \right) x_{t-1} + \frac{C(x_t, x_0)}{\text{与 } x_{t-1} \text{ 无关}}}_{x_{t-1} \text{ 均值}} \right) \right) \quad (17)$$

与高斯分布的概率密度函数相对应，可得

$$\begin{aligned} q(x_{t-1} | x_t, x_0) &= N(x_{t-1}; \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t I) \\ \tilde{\mu}_t(x_t, x_0) &= \frac{\sqrt{\alpha_{t-1}} \beta_t}{1-\alpha_t} x_0 + \frac{\sqrt{\alpha_t} (1-\alpha_{t-1})}{1-\alpha_t} x_t \\ \tilde{\beta}_t &= \frac{1-\alpha_{t-1}}{1-\alpha_t} \beta_t \end{aligned} \quad (18)$$

附录 2 L_{simple} 损失项

对于模型预测的均值 $\mu_\theta(x_t, t)$ ，根据多元高斯分布的 KL 散度求解方法可得

$$L_{t-1} = \mathbb{E}_q \left[\frac{1}{2\sigma_t^2} \|\tilde{\mu}_t(x_t, x_0) - \mu_\theta(x_t, t)\|^2 \right] + C \quad (19)$$

其中， C 是和模型参数无关的常量。将高斯分布 $q(x_t | x_0) \sim N(x_t; \sqrt{\alpha_t} x_0, (1-\alpha_t)I)$ 和 μ_θ 分别进行参数化，代入式(19)可得

$$L_{t-1} - C = \mathbb{E}_{x_0, \epsilon} \left[\frac{\beta_t^2}{2\sigma_t^2 \alpha_t (1-\alpha_t)} \left\| \epsilon - \epsilon_\theta \left(\sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t} \epsilon, t \right) \right\|^2 \right] \quad (20)$$

其中， ϵ_θ 是根据 x_t 预测 ϵ 的函数。反向过程对 $x_{t-1} \sim p_\theta(x_{t-1} | x_t)$ 的采样变为计算

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z \quad (21)$$

其中， z 是标准高斯噪声。由此可定义简化的训练目标 L_{simple} ，使模型更容易实现，实践中发现也有利于提高样本质量，表示为

$$L_{\text{simple}}(\theta) := \mathbb{E}_{t, x_0, \epsilon} \left[\left\| \epsilon - \epsilon_\theta \left(\sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t} \epsilon, t \right) \right\|^2 \right] \quad (22)$$

参考文献：

[1] 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.
ZHAO Z H, TONG Y F, CUI J Y. Device-independent Wi-Fi fingerprinting indoor localization model based on domain adaptation[J]. Journal on Communications, 2022, 43(4): 143-153.

[2] 周牧, 王斌, 田增山, 等. 室内 BLE/MEMS 跨楼层融合定位算法[J].

通信学报, 2017, 38(5): 1-10.

ZHOU M, WANG B, TIAN Z S, et al. Indoor BLE and MEMS based multi-floor fusion positioning algorithm[J]. Journal on Communications, 2017, 38(5): 1-10.

[3] WANG R R, LI Z H, LUO H Y, et al. A robust Wi-Fi fingerprint positioning algorithm using stacked denoising autoencoder and multi-layer perceptron[J]. Remote Sensing, 2019, 11(11): 1293.

[4] AI H J, SUN X, TAO J J, et al. DRVAT: exploring RSSI series representation and attention model for indoor positioning[J]. International Journal of Intelligent Systems, 2022, 37(7): 4065-4091.

[5] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 6000-6010.

[6] TALVITIE J, RENFORS M, LOHAN E S. Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals[J]. IEEE Transactions on Vehicular Technology, 2015, 64(4): 1340-1353.

[7] SINHA R S, LEE S M, RIM M, et al. Data augmentation schemes for deep learning in an indoor positioning application[J]. Electronics, 2019, 8(5): 554.

[8] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.

[9] NJIMA W, CHAFII M, CHORTI A, et al. Indoor localization using data augmentation via selective generative adversarial networks[J]. IEEE Access, 2021, 9: 98337-98347.

[10] LAN T, WANG X M, CHEN Z K, et al. Fingerprint augment based on super-resolution for Wi-Fi fingerprint based indoor localization[J]. IEEE Sensors Journal, 2022, 22(12): 12152-12162.

[11] SUGASAKI M, SHIMOSAKA M. Robustifying Wi-Fi localization by between-location data augmentation[J]. IEEE Sensors Journal, 2022, 22(6): 5407-5416.

[12] KINGMA D P, WELLMING M. Auto-encoding variational Bayes[J]. arXiv Preprint, arXiv: 1312.6114, 2013.

[13] SOHL-DICKSTEIN J, WEISS E A, MAHESWARANATHAN N, et al. Deep unsupervised learning using nonequilibrium thermo dynamics[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning. New York: ACM Press, 2015: 2256-2265.

[14] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM Press, 2020: 6840-6851.

[15] DHARIWAL P, NICHOL A. Diffusion models beat GANs on image synthesis[J]. arXiv Preprint, arXiv: 2105.05233, 2021.

[16] HO J, SALIMANS T. Classifier-free diffusion guidance[J]. arXiv Preprint, arXiv: 2207.12598, 2022.

[17] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 10674-10685.

[18] NICHOL A, DHARIWAL P, RAMESH A, et al. GLIDE: towards photorealistic image generation and editing with text-guided diffusion models[J]. arXiv Preprint, arXiv: 2112.10741, 2021.

- [19] CHEN K M, CHANG R Y. A comparative study of deep-learning-based semi-supervised device-free indoor localization[C]//Proceedings of IEEE Global Communications Conference (GLOBECOM). Piscataway: IEEE Press, 2022: 1-6.
- [20] KUMAR S, HEGDE R M, TRIGONI N. Gaussian process regression for fingerprinting based localization[J]. Ad Hoc Networks, 2016, 51: 1-10.
- [21] ZOU H, CHEN C L, LI M X, et al. Adversarial learning-enabled automatic Wi-Fi indoor radio map construction and adaptation with mobile robot[J]. IEEE Internet of Things Journal, 2020, 7(8): 6946-6954.
- [22] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002, 16: 321-357.
- [23] INOUE H. Data augmentation by pairing samples for images classification[J]. arXiv Preprint, arXiv: 1801.02929, 2018.
- [24] ZHANG H, Cisse M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[J]. arXiv Preprint, arXiv: 1710.09412, 2017.
- [25] NICHOL A, DHARIWAL P. Improved denoising diffusion probabilistic models[J]. arXiv Preprint, arXiv: 2102.09672, 2021.
- [26] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 5967-5976.
- [27] HENIGHAN T, KAPLAN J, KATZ M, et al. Scaling laws for autoregressive generative modeling[J]. arXiv Preprint, arXiv: 2010.14701, 2020.

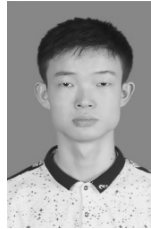
[作者简介]



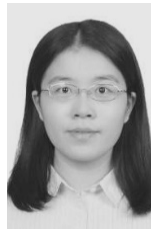
艾浩军(1972-),男,湖北汉川人,博士,武汉大学副教授、博士生导师,主要研究方向为普适计算、计算机视觉、无线感知等。



曾维珂(2000-),女,江西南昌人,武汉大学硕士生,主要研究方向为室内定位、计算机视觉。



陶荆杰(1999-),男,江西南昌人,武汉大学硕士生,主要研究方向为普适计算、机器学习。



徐锦盈(1999-),女,湖北武汉人,武汉大学硕士生,主要研究方向为室内定位、计算机视觉。



常含笑(2000-),男,河南济源人,武汉大学硕士生,主要研究方向为机器学习、计算机视觉。