

引用本文:马芊慧,王旭东,吴楠,等.基于注意力机制WGAN-GP的可见光指纹定位方法[J].光通信技术,2026,50(1):1-9.

基于注意力机制WGAN-GP的可见光指纹定位方法

马芊慧,王旭东*,吴楠,曲佳

(大连海事大学 信息科学技术学院,辽宁 大连 116026)

摘要:为减少可见光指纹定位在指纹建库中的人力物力消耗并提升传统数据增强算法的全局感知能力,提出基于注意力机制Wasserstein GAN-梯度惩罚(GP)项的可见光指纹定位方法。该方法在生成器和判别器中嵌入多头自注意力模块,对指纹空间分布特征进行增强,并引入GP项以解决模式坍塌问题;采用发光二极管(LED)的光强信号作为特征构建指纹数据库,利用增强后数据训练轻量化网络实现位置估计。仿真实验表明:所提模型生成数据质量较传统GAN显著提高;在5 m×5 m×3 m的室内场景下,实测定位误差可达6~8 cm。

关键词:室内可见光定位;指纹数据生成;注意力机制;接收信号强度;生成对抗网络

中图分类号:TN929.1 **文献标志码:**A **文章编号:**1002-5561(2026)01-0001-09

DOI:10.13921/j.cnki.issn1002-5561.2026.01.001

Visible light fingerprint localization method based on WGAN-GP with attention mechanism

MA Qianhui, WANG Xudong*, WU Nan, QU Jia

(Information Science Technology College, Dalian Maritime University, Dalian Liaoning 116026, China)

Abstract: To reduce the labor and material consumption of visible light fingerprint localization in fingerprint database construction and enhance the global perception ability of traditional data augmentation algorithms, a visible light fingerprint localization method based on the attention mechanism Wasserstein GAN-Gradient Penalty (GP) term is proposed. This method embeds multi-head self-attention modules in the generator and discriminator to enhance the spatial distribution characteristics of fingerprints, and introduces GP to solve the mode collapse problem. The light intensity signal of a light-emitting diode (LED) is used as a feature to construct a fingerprint database, and the enhanced data is used to train a lightweight network for position estimation. Simulation experiments show that the quality of data generated by the proposed model is significantly improved compared to traditional GANs. In an indoor scene of 5 m×5 m×3 m, the measured positioning error can reach 6~8 cm.

Key words: indoor visible light positioning, fingerprint data generation, attention mechanism, received signal strength, generative adversarial network

0 引言

随着全球智能设备与物联网的快速发展,定位与导航已成为各类应用不可或缺的基础服务^[1]。以射频识别、蓝牙定位、超宽带定位、Wi-Fi定位及计算机视觉

定位等为代表的传统室内定位技术,通常存在定位精度低、部署成本高、计算资源消耗大以及易受电磁干扰等问题^[2-6]。可见光定位(VLP)技术兼具高精度、低能耗、抗电磁干扰等优点,被认为是未来室内定位的重要发展方向^[7]。在诸多VLP方法中,指纹定位法以其定位精度高、成本低廉、适用场景广泛等优势得到广泛关注。可见光指纹定位的工作原理是:在预设的定位区域内,于不同参考点上采集诸如接收信号强度(RSS)、到达时间差(TDOA)或者波达方向(AOA)等可见光信号特征,并以此建立位置指纹数据库。定位时,通过将实测信号特征与数据库中的指纹进行匹配,即可估计用户位置。然而,可见光指纹定位法需要在部署区域内大量采集光信号特征参数以构建指

收稿日期:2025-09-20。

基金项目:国家自然科学基金项目(61801074)资助。

作者简介:马芊慧(2001—),女,黑龙江哈尔滨市人,硕士研究生,本科毕业于燕山大学通信工程专业,现就读于大连海事大学信息科学技术学院信息与通信工程专业,主要研究方向为可见光通信、室内可见光定位。曾荣获大连海事大学研究生学业三等奖学金。

*通信作者:王旭东(1967—),男,博士,教授,硕士生导师,主要从事光无线通信调制技术、可见光无线通信及定位、基于人工智能的无线通信物理层设计等方面的研究工作。



马芊慧,王旭东,吴楠,等:基于注意力机制 WGAN-GP 的可见光指纹定位方法

纹数据库,这一过程耗时耗力。此外,该方法还易受灯具老化、环境变化等因素影响,导致定位精度下降,从而制约了其大规模应用^[8]。为解决实际指纹采集成本高、数据易失效等问题,近年来研究者提出了通过深度生成模型合成指纹数据的方法,旨在提高指纹库的丰富度、稳定性及泛化能力,从而提升定位模型性能并减少现场数据采集需求。文献[9]提出了一种基于生成对抗网络(GAN)的高精度室内定位方法,利用等间距网格化采集的RSS信号构建低密度初始指纹数据库,通过GAN对稀疏采样数据进行特征学习与空间分布建模,有效解决了传统方法依赖大规模实测数据的问题,但是该方法对模型设计与训练过程中的模式坍塌和数据多样性等问题考虑不足。文献[10]提出了一种基于GAN与卷积神经网络(CNN)联合的定位方案,通过引入参考点序号作为条件信息,结合最小二乘损失函数优化生成对抗网络,有效生成信道状态信息(CSI)指纹数据,缓解传统方法对海量实测样本的依赖,但其定位精度有所下降。文献[11]提出了一种基于条件稀疏自编码生成对抗网络的室内定位方法(cGAN-SAE)。该方法通过增加自编码器的隐藏层与输出层来强化特征提取能力,从而引导生成器学习并生成指纹数据的关键特征。所生成的数据用于训练一个卷积长短时记忆网络模型,以进行在线定位。数据集上的预测及实际应用测试均验证了其定位性能的有效提升,但该方法在训练与部署阶段的计算代价较高。文献[12]则提出了一种基于改进 ShuffleNet V2 网络的室内可见光指纹定位算法。该算法通过减少部分卷积层并增加全连接层,旨在缓解梯度爆炸与过拟合风险,同时提升定位的实时性与整体效率。尽管其轻量化网络设计有效提高了定位效率,但针对指纹库构建过程本身的高昂代价,该研究尚未提出有效的优化方案。

为进一步解决传统指纹库构建过程中存在的人力物力消耗及传统GAN中缺乏全局感知能力等问题,本文提出基于注意力机制 Wasserstein GAN-梯度惩罚(GP)项的可见光指纹定位方法(下文简称 AT-WGAN-GP)。

1 定位模型

本文在一个 5 m×5 m×3 m 的室内场景下建立 VLP 系统模型,如图1所示。4个发光二极管(LED)以正方形拓扑结构安装于天花板,其平面坐标分别为(1.25 m, 1.25 m), (1.25 m, 3.75 m), (3.75 m, 1.25 m), (3.75 m,

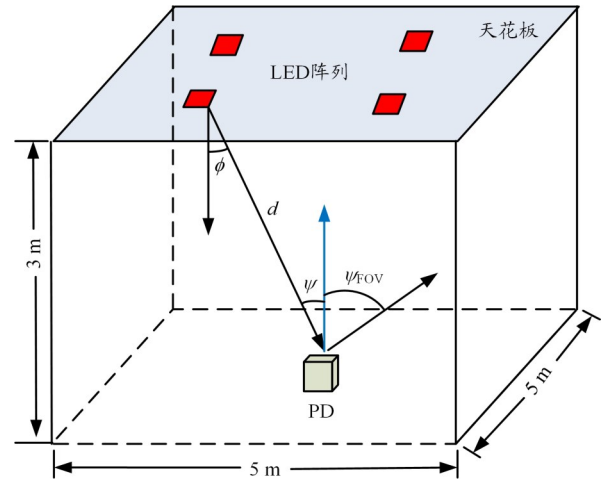


图1 VLP系统模型

3.75 m),高度均为H。

鉴于直射路径信号在可见光通信(VLC)系统中占主导地位^[13],本文主要考虑直射链路(LOS)信道模型。根据Lambertian辐射模型^[14],其信道直流增益可表示为

$$H(0) = \frac{(m+1)}{2\pi d^2} A \cos^m \phi T_s(\psi) g(\psi) \cos \psi \quad (1)$$

式中: m 为朗伯系数, $m = -\ln 2 / \ln(\cos \phi_{1/2})$, $\phi_{1/2}$ 是LED的半功率角。 d 为发射端LED到接收端光电探测器(PD)的距离, A 为PD的接收面积, ϕ 为发射端的辐射角, ψ 为接收端的入射角, $g(\psi)$ 为聚光器增益, $T_s(\psi)$ 为光滤波器增益, ψ_{FOV} 为聚光器的视场角(FOV), $0 \leq \psi \leq \psi_{FOV}$ 。

设LED发出光信号的功率为 P_t ,则通过LOS信道到达接收端的信号功率 P_r 可表示为

$$P_r = H(0)P_t \quad (2)$$

该模型用于仿真生成不同位置点的接收信号强度,以此构建初始指纹数据库。

2 AT-WGAN-GP 的指纹生成方法

为了有效生成具有真实空间分布特性的指纹数据,需要构建一套能够同时捕捉局部细节与全局结构的生成模型。AT-WGAN-GP旨在结合生成对抗框架的分布建模能力与注意力机制的全局关联建模优势,从而在低采样条件下仍能够生成高质量指纹数据。

2.1 AT-WGAN-GP 网络

WGAN在传统GAN基础上引入Wasserstein距离,替代原有的KL散度或JS散度,更有效衡量概率分布差异,缓解了梯度消失与模式塌陷问题。然而,WGAN依赖权重裁剪以满足1-Lipschitz约束,限制了判别器

的表达能力,难以捕捉复杂数据特征。WGAN-GP则通过引入GP替代权重裁剪,避免了参数趋于边界值的问题,提升了判别器性能,但对全局特征的感知能力仍显不足。为此,本文在生成器与判别器中引入注意力模块,以显著增强模型对数据全局空间特征的学习能力。

2.1.1 注意力机制模块

注意力机制^[15]是一种参照人类认知注意力原理设计的机器学习模型。其核心思想是为输入信息的不同部分赋予动态权重,通过构建注意力概率分布来强调关键特征、抑制次要信息,从而实现对关键输入的聚焦。本文中的注意力机制模块是Transformer中编码层的变体,其示意图如图2所示。该模块主要包含4个模块:层归一化、多头注意力机制、前馈网络和双重残差连接。

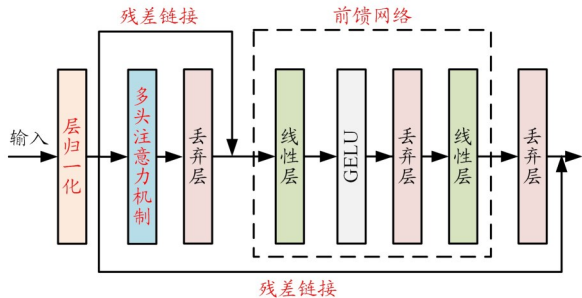


图2 注意力机制模块

层归一化^[16]:该模块将输入数据转化成均值为0、标准差为1的分布,使预处理的数据保持在一个稳定的范围内,防止了过大或过小导致的梯度消失或梯度爆炸问题。

多头注意力机制^[17]:该模块通过权重值反映对不同信息的关注程度,可视为查询矩阵(Query)、键(Key)以及加权平均值(Value)。该机制通过并行多个独立K-V来增强模型表征能力,其核心思想是将输入特征投影到多个子空间进行差异性特征学习,每个注意力头独立计算Q、K、V的关联权重并生成子特征,最终将各头的输出拼接形成综合特征权重。在生成VLP指纹的任务中,多头注意力机制学习各接入点(AP)间的

指纹特征,能够使用不同序列位置的不同子空间的表征信息来进行序列数据处理。该机制原理如图3所示。

前馈网络:首先,该网络第一个线性层将输入特征维度扩大4倍,以增强模型的表达能力;随后,通过高斯误差线性单元^[18](GELU)激活函数。为了对生成的高阶特征进行正则化以防止过拟合,此处设置了10%的丢弃层。最后,第二个线性层将特征维度压缩回原始尺寸,以确保其输出维度与输入维度一致,从而便于与该模块的输入进行残差连接。GELU表达式为

$$G_{\text{GELU}}(y) = 0.5y \left\{ 1 + \tanh \left[\sqrt{\frac{2}{\pi}} (y + 0.047715y^3) \right] \right\} = y \times S_{\sigma}(1.702y) \quad (3)$$

式中: y 为激活函数的输入,即神经网络中某一神经元的线性加权和; S_{σ} 为Sigmoid类的高斯累积分布函数。

双重残差连接:输入数据层归一化后分别与多头注意力层和前馈层连接残差路径,其输出在两级结构中均与对应的输入通过跳跃连接相加。第一级残差用于保持多头注意力提取过程中的原始特征,第二级残差用于在前馈非线性映射后进一步强化信息保留。两级残差在相加前均应用Dropout结构随机失活10%神经元,防止过拟合。

2.1.2 WGAN-GP 生成器设计

WGAN-GP生成器主要由3个部分组成:输入处理层、特征变换层、输出层,如图4所示。

输入处理层:首先,通过一个线性层,将随机生成的低维噪声向量转换并提升到128维特征空间,以增加模型的表征能力。随后,进行层归一化,用于稳定各层的输入分布,缓解深层网络训练中常见的梯度问题,加速模型收敛。再进行LeakyReLU激活函数为网

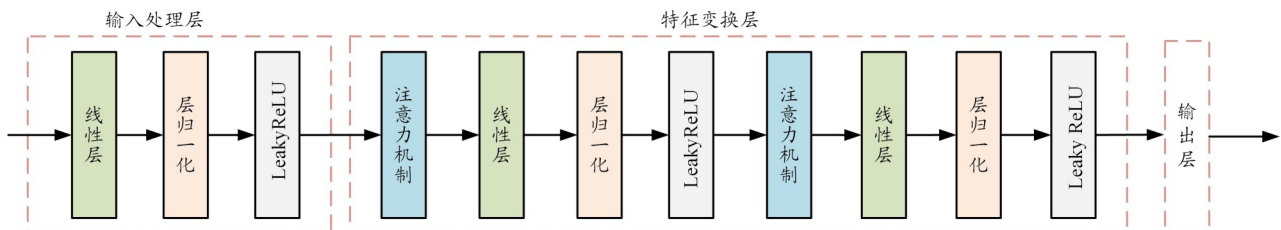


图4 WGAN-GP生成器结构

马芊慧,王旭东,吴楠,等:基于注意力机制 WGAN-GP 的可见光指纹定位方法

络引入非线性表达能力,LeakyReLU 为其负区间设定了 20% 的“渗漏”斜率,其激活函数表达式为

$$R_{\text{LeakyReLU}}(y) = \max(0, y) + \alpha \times \min(0, y) \quad (4)$$

式中: α 为 LeakyReLU 在负区间的斜率系数,用于控制负值通过激活函数时的衰减程度。

特征变换层:首先,引入一个 2 层的 128 维注意力模块,通过自注意力机制深入地挖掘高维特征之间的复杂依赖关系,以增强生成数据的全局结构一致性。随后,使用一个线性层将 128 维特征压缩至 64 维,依次经过层归一化和 LeakyReLU 激活函数。之后,通过一个 1 层的 64 维注意力模块,在降维后的特征空间中进行浅层的依赖关系处理。最后,再经一个线性层将特征进一步压缩至 32 维,并通过层归一化和 LeakyReLU 激活函数,输出最终变换后的特征。

输出层:直接线性映射到 6 维目标维度。

WGAN-GP 生成器采用 Adam 优化器进行训练,学习率设为 0.000 1。优化目标是使判别器将其生成的样本判定为“真实”,即最大化判别器对生成样本的输出值。由于优化器默认最小化损失,所以将生成器的损失函数 L_G 定义为

$$L_G = -E_{x \sim p_g(x)}[D(G(z))] \quad (5)$$

式中: x 为真实数据样本, $p_g(x)$ 为生成器生成的数据分布, $E[\cdot]$ 为求期望运算, $G(z)$ 为生成器对噪声映射后得到的生成样本, $D(\cdot)$ 为判别器网络,用于估计输入样本属于真实数据分布的概率。

2.1.3 WGAN-GP 判别器设计

WGAN-GP 判别器的输入为真实数据或生成数据,其功能是对输入的真伪进行判断并输出概率。判别器结构如图 5 所示,主要由频谱归一化^[19]、Leaky-ReLU 激活函数和注意力机制模块构成。线性输入层首先将 6 维数据映射至 32 维特征空间,随后通过频谱归一化约束权重矩阵的谱范数,以限制 Lipschitz 常数,从而稳定对抗训练。对于权重矩阵 $W \in R^{m \times n}$,其谱范数 $\sigma(W)$ 定义为

$$\sigma(W) = \max_{x \in R^n, x \neq 0} \frac{\|Wx\|_2}{\|x\|_2} \quad (6)$$

式中: R 表示实数域, R^n 表示 n 维实向量空间, R^m 表示 m

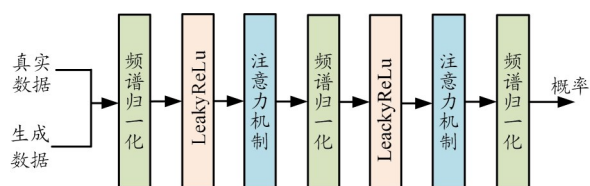


图5 WGAN-GP 判别器结构

维实向量空间。频谱归一化将 W 除以 $\sigma(W)$,得到归一化后的权重 W_{SN} 为

$$W_{\text{SN}} = \frac{W}{\sigma(W)} \quad (7)$$

判别器中的 LeakyReLU 激活函数用于引入非线性并缓解梯度消失问题。注意力模块则负责捕捉 32 维特征之间的全局依赖关系,通过动态分配权重,使判别器能更聚焦于关键特征。随后,特征经过频谱归一化处理,并由线性层从 32 维扩展至 64 维,这一过程显著提升了判别器的特征表达能力。然后,一个深层的 64 维注意力模块进一步提取更复杂的特征依赖关系。最终,输出层将 64 维高级特征映射为 1 维标量,该标量的值越大,代表判别器判定输入数据来自真实分布的概率越高。

判别器采用学习率为 0.000 4 的 Adam 优化器进行训练。其核心在于引入 GP 机制:对真实数据与生成数据之间的插值样本计算梯度,并按式(8)构造判别器的损失函数 L_D ,其中 GP 的权重作为超参数被设为 10。

$$L_D = E_{x \sim p_g(x)}[D(x)] - E_{x \sim p_r(x)}[D(x)] + \lambda E_{\hat{x}} \left\{ \left[\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right]^2 \right\} \quad (8)$$

式中: $D(x)$ 为将样本判别为真样本的概率分布, $p_r(x)$ 为真实数据分布, $\lambda E_{\hat{x}} \left\{ \left[\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right]^2 \right\}$ 为 GP, λ 为 GP 项的权重系数,用于调节 Lipschitz 约束强度。

2.2 指纹生成算法

本文提出的 AT-WGAN-GP 模型包括指纹库构建、数据增强、定位模型训练和在线预测 4 个阶段,算法流程图如图 6 所示。

指纹库构建过程如下:首先,将接收平面均匀划分为 N 个网格,中心坐标作为标签位置, $(x_i, y_i), i=1, 2, \dots, N$,表示 N 个网格中第 i 个网格的中心坐标。随后,在每个网格点上采集来自 4 个 LED 的 RSS,形成对应的 RSS 特征向量 $P_i=[P_{i1}, P_{i2}, P_{i3}, P_{i4}]$ 。每个位置坐标与其 RSS 向量共同构成一条指纹数据,由此建立的指纹库数据库如表 1 所示。所构建的指纹库作为原始数据集,将被输入到 AT-WGAN-GP 网络中以生成新的指纹数据,用于后续的数据增强阶段。

为了更有效地捕捉每个参考点上的特征信息,本实验在一个尺寸为 5 m×5 m×3 m 的室内场景中进行。在接收平面(高度 $h=0.8$ m)上,按 0.5 m 的间隔均匀划分出 121 个参考点。在每个参考点上,采集来自 4 路 LED 的接收信号强度,共获取 100 组 RSS 数据,总计形

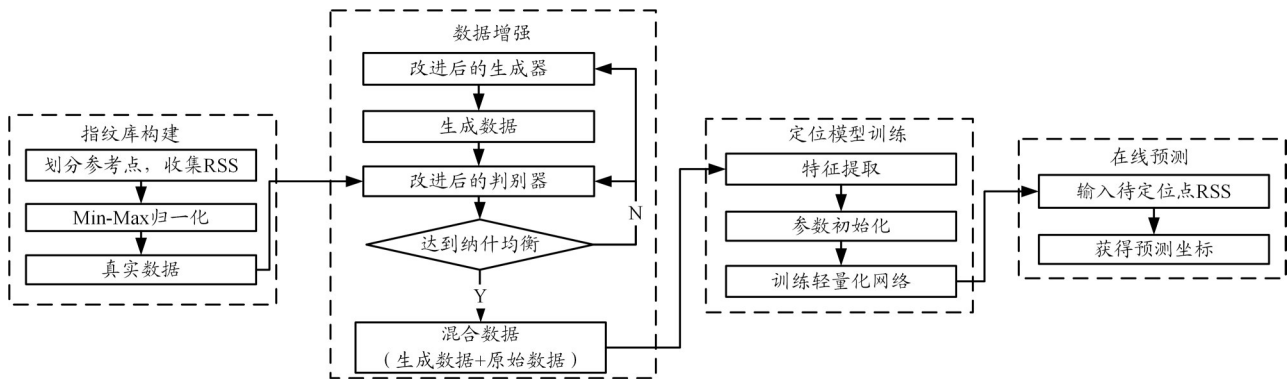


图6 AT-WGAN-GP算法流程

表1 指纹数据库

参考点	坐标	RSS			
1	(x_1, y_1)	P_{11}	P_{12}	P_{13}	P_{14}
2	(x_2, y_2)	P_{21}	P_{22}	P_{23}	P_{24}
3	(x_3, y_3)	P_{31}	P_{32}	P_{33}	P_{34}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	(x_N, y_N)	P_{N1}	P_{N2}	P_{N3}	P_{N4}

成12 100组原始数据。随后,将全部数据随机划分为训练集与测试集,其中90%用于训练,10%用于测试。为提升神经网络训练的效率和收敛速度,在输入网络前,对原始RSS数据采用最大最小值归一化方法进行了预处理。

针对包含 N 组测量样本的数据集 $\{x_{i,n}\}_{i=1}^N$,通过特征缩放实现归一化处理:

$$\hat{x}_{i,n} = \frac{x_{i,n} - \min(x_n)}{\max(x_n) - \min(x_n)} \quad (9)$$

采用最大最小值标准化将各特征值归一化至[0, 1]区间,可有效统一量纲、消除特征间的耦合效应,从而提高模型训练的数值稳定性。该处理有助于避免WGAN-GP判别器在训练中出现梯度剧烈波动,保障其平稳收敛。随后,将原始指纹数据与随机噪声一同输入到AT-WGAN-GP网络中,以生成新的指纹数据。在此过程中,注意力机制被用于抑制噪声干扰并强化关键的指纹特征表达。生成的数据将与原始数据混合,共同用于训练2个轻量化网络(MobileNet V2和ShuffleNet V2),从而得到训练完成的定位模型。在定位阶段,将待测点采集的RSS信号输入至已训练的模型中,即可输出其对应的预测坐标,完成位置估计。

2.3 指纹生成质量评估

为客观量化AT-WGAN-GP模型所生成指纹数据

的质量,并验证其能否有效逼近真实数据的分布,本文引入多维度评估指标,对生成数据进行全面评估。

1) t-SNE^[20] (t-Stochastic Neighbor Embedding) 可视化

t-SNE是一种非线性降维技术,广泛用于高维数据的可视化,尤其在机器学习中用于探索数据分布或评估生成模型的输出质量。将高维数据非线性映射到低维空间,同时保留数据点之间的局部相似性结构。给定高维空间中数据点 x_i 和 x_j ,条件概率 p_{ji} 表示在 x_j 的邻域内,点 x_i 被选为邻居的似然度,其定义为

$$p_{ji} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)} \quad (10)$$

数据标准差 σ_i 通过困惑度调节,控制局部邻域的大小。 N_{total} 为数据总数,对称后的联合概率为

$$p_{ij} = \frac{p_{ji} + p_{ij}}{2N_{\text{total}}} \quad (11)$$

在生成模型评估中,将真实数据和生成数据合并为高维数据集,使用t-SNE降维至二维,可视化两类数据的分布重叠情况。若生成数据点与真实数据点混合均匀,说明生成质量高;若生成数据形成孤立簇,可能存在模式崩溃。

2) 弗雷歇起始距离(FID)

FID是衡量生成对抗网络生成数据与真实数据分布相似性的经典指标,它通过特征空间中的分布差异来量化生成质量,更关注生成数据的统计特性而非分类特性。FID的目标是计算生成数据分布 $p_g(x)$ 和真实数据分布 $p_r(x)$ 在特征空间中的FID,值越小代表生成效果越好,其公式可表示为

$$F_{\text{FID}}(u_r, u_g; C_r, C_g) = \|u_r - u_g\|_2^2 + T_r(C_r + C_g - \sqrt{C_r C_g}) \quad (12)$$

马芊慧,王旭东,吴楠,等:基于注意力机制WGAN-GP的可见光指纹定位方法

式中: u_g 和 C_g 分别表示生成数据的均值和协方差, u_r 和 C_r 分别表示真实数据的均值和协方差, T_r 表示真实数据矩阵的迹。FID直接量化分布相似性,对分类器不敏感,计算高效且稳定,适合大规模数据集,广泛应用于GAN、变分自编码器(VAE)等生成模型的评估。

3)平均绝对偏差(MAD)

MAD是一种用于量化数值序列平滑性的统计指标,通过计算相邻数据点之间绝对值的平均值,评估序列的局部波动程度。MAD值越高表明相邻点波动剧烈,序列越不稳定,反之越稳定。 k 表示序列中元素个数, L_i 为序列中第 i 个元素,MAD公式表示为

$$M_{\text{MAD}} = \frac{1}{k-1} \sum_{i=1}^{k-1} |L_{i+1} - L_i| \quad (13)$$

4)GP项

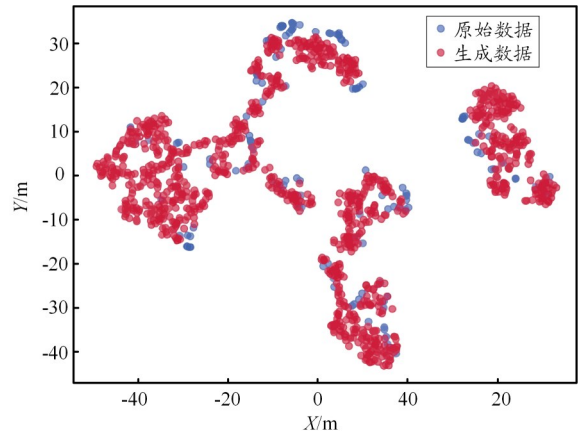
传统WGAN通过权重剪裁强制Lipschitz约束,但可能导致梯度消失或爆炸。WGAN-GP则通过GP直接约束判别器的梯度范数,使其在插值样本点附近接近1,从而更稳定地满足Lipschitz条件。在真实数据与生成数据的随机插值点 $\hat{x} = \varepsilon x + (1 - \varepsilon)G(z)$ 处,计算判别器输出的梯度范数 $\|\nabla_x D(\hat{x})\|_2$,并通过GP项对其进行约束。该GP项旨在迫使判别器在插值点附近的梯度范数稳定在1左右,避免局部梯度过大或过小,从而从统计意义上保证Lipschitz约束的全局有效性。其中, ε 为随机采样的插值系数,用于在真实样本与生成样本之间构造随机插值点。

当生成数据与真实数据差异较大时,二者分布重叠区域较小,导致随机插值点 $\hat{x}$ 更偏向真实或生成数据分布的一侧。此时,判别器在这些点的梯度范数偏离1,使GP值偏大。随着生成分布不断逼近真实分布,插值区域趋于真实分布,判别器梯度趋于平缓,其范数接近1,GP值也随之减小。因此,GP值的大小反映了判别器对分布差异的敏感程度,GP值低表明生成分布已接近真实分布,若GP值趋近0则表示模型趋于收敛;反之,GP值较高则表示生成数据与真实数据之间仍存在明显差距。

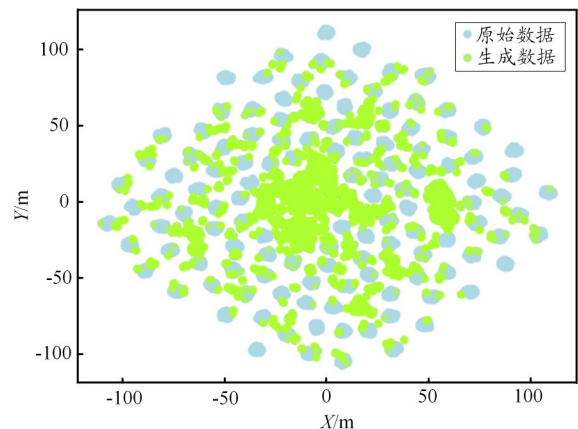
3 实验结果及分析

3.1 生成数据质量评估

本文从数据分布可视化、量化评估指标和训练过程收敛这3个维度,对比分析传统GAN与AT-WGAN-GP网络的生成数据质量,得到的t-SNE可视化对比结果如图7所示。可以看出,AT-WGAN-GP网络生成数



(a) AT-WGAN-GP



(b) GAN

图7 不同网络的t-SNE可视化对比

据与原始数据在特征空间中的分布呈现高度重叠,说明生成数据的质量较高;而传统GAN的生成数据样本则聚集在局部区域,且与原始数据分布存在明显差异。

设训练步长为15 000,此时不同网络性能对比如表2所示。可以看出,传统GAN的FID值为8.42,而AT-WGAN-GP降至5.13,相对提升了39.07%,验证了模型结构的有效性。在收敛平滑性方面,AT-WGAN-GP的MAD指标较传统GAN降低了76.18%,这得益于GP对参数更新的正则化作用。

表2 不同网络性能对比

网络	FID	MAD
GAN	8.42	0.048 7
AT-WGAN-GP	5.13	0.011 6

不同模型生成器与判别器的损失函数曲线^[21]如图8所示。可以看出,传统GAN模型生成器和判别器收敛平滑性较差(印证了表2中数据)。而AT-WGAN-

GP模型生成器和判别器收敛速度更快,在训练步长为1 000时已呈现明显的收敛趋势,模型的生成能力更强。AT-WGAN-GP模型的GP项损失函数的收敛曲线如图9所示。可以看出,GP项损失函数值最终稳定在 0.005 ± 0.001 的区间内。该值接近于0,满足梯度约束条件,确保了生成样本的分布保真度。

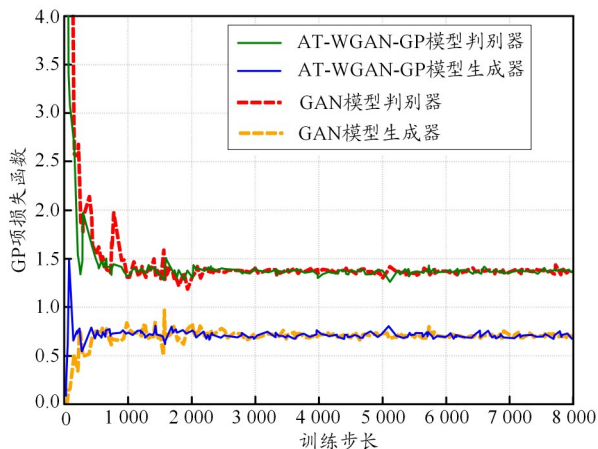


图8 不同模型生成器与判别器损失函数图

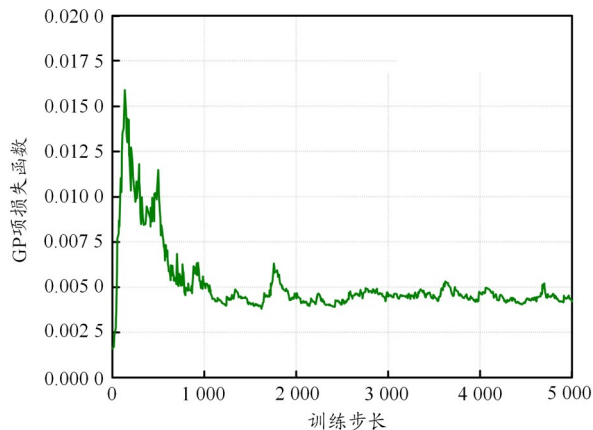


图9 AT-WGAN-GP模型的GP项收敛曲线

3.2 不同生成策略定位性能

为了解决室内定位中指纹数据不足的问题,本文提出基于位置标签的3种不同的生成策略^[22]:原点扩充(PE)、等间隔扩充(SE)、随机扩充(RE)。原始数据采集点的间隔均匀为0.5 m。PE策略固定以原始数据采集点坐标作为标签,使用AT-WGAN-GP模型生成各标签对应的RSS指纹数据。SE策略引入了2个超参数:间隔 u 和倍数 W ,本文选取固定间隔 $u=0.2$ m作为一个位置标签。RE策略则不设置任何标签,直接在 $5\text{ m} \times 5\text{ m}$ 的平面内随机生成指纹数据。

将训练集输入到AT-WGAN-GP模型中,PE策略在每隔0.5 m间隔的点上扩充50组数据,共得到 $121 \times$

50组RSS数据;SE取 $W=10$ 和 $u=0.2$ m,生成 676×10 组RSS数据;RE则生成6 500组RSS数据。3种扩充方式生成的数据量基本一致。将得到的生成数据输入到MobileNet V2中进行定位,得到的累积概率分布函数(CDF)曲线如图10所示。计算得出,SE、PE、RE的平均定位误差分别为10.35、13.49、15.65 cm。可以看出,SE的CDF曲线左移明显,表明其定位精度更高;PE和RE在大约0.2 m误差前的曲线几乎重合,但RE策略具有较强的不确定性,其CDF曲线存在拖尾现象,部分定位点误差较大。SE策略与常规指纹库构建方式相似度很高,超参数 u 和 W 可灵活设置,能满足对不同房间大小和不同扩充倍数的需求。

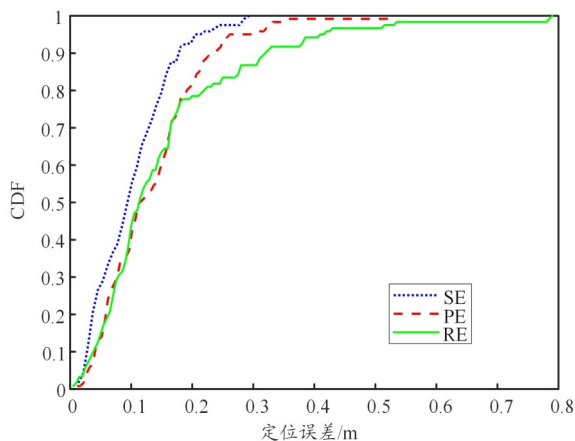


图10 不同生成策略CDF曲线

3.3 不同生成数据量定位性能

为探究生成数据量对定位性能的影响,本文控制SE策略中的扩增倍数 W 分别取10、20、30、40,并将生成数据 G 、生成数据 G +原始数据 R 以及原始数据 R 分别输入到MobileNet V2网络中进行定位,得到性能结果如图11所示。可以看出,仅使用原始数据 R 训练时,由于数据量有限(121×80 组),模型训练不充分,定位误差为12.80 cm。随着扩增倍数 W 增大,仅使用生成数据 G 进行定位时,性能较原始数据基准分别提升了19.14%、28.05%、34.14%和35.55%。当 W 从30增至40时,定位精度提升幅度仅为1.41%,考虑到计算成本与性能增益的边际效益,实际应用中建议选取 $W=30$ 。进一步地,使用生成数据与原始数据混合集 $G+R$ 训练时,相较于仅使用生成数据,定位精度平均提升了6.93%。但随着 W 增大,原始数据对混合训练的性能贡献逐渐减弱,说明AT-WGAN-GP模型生成的数据已具备较高的独立表征能力,从侧面验证了所生成指纹数据具有良好的性能。

马芊慧,王旭东,吴楠,等:基于注意力机制WGAN-GP的可见光指纹定位方法

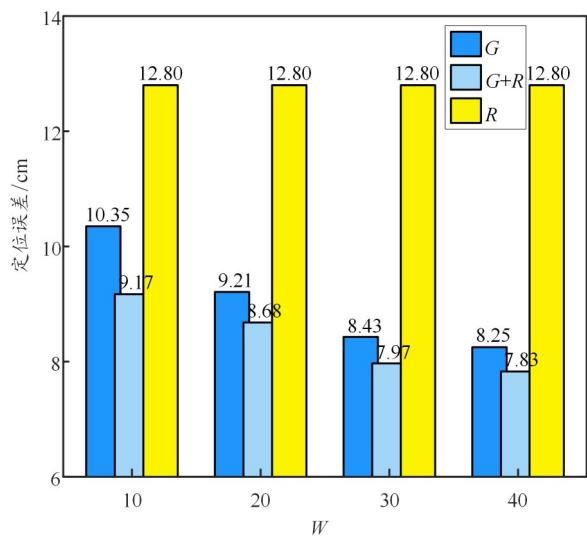


图11 不同生成数据量定位性能

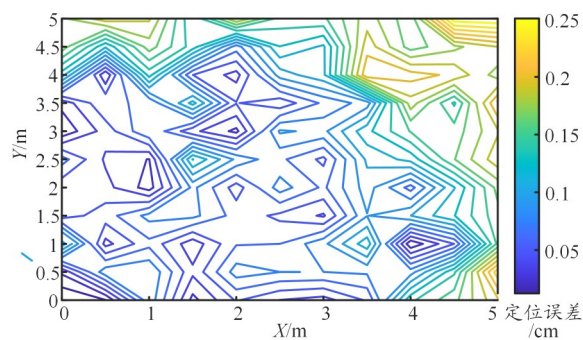
3.4 注意力机制消融实验定位对比

将图4和图5中的注意力机制用线性层代替,其余结构与参数保持不变。将121×80组原始数据分别输入WGAN-GP与AT-WGAN-GP,以等间距 $u=0.2$ m生成RSS指纹数据(每点30组)。随后,将生成数据与原始数据混合,分别用于训练MobileNet V2和ShuffleNet V2定位网络。最终在测试集上的定位性能对比如表3所示。由表3可知,AT-WGAN-GP生成数据将MobileNet V2和ShuffleNet V2的平均定位精度分别提升了21.40%和28.50%。这验证了注意力机制通过提高生成数据质量来降低定位误差的有效性,且其对ShuffleNet V2网络的提升效果更为突出。

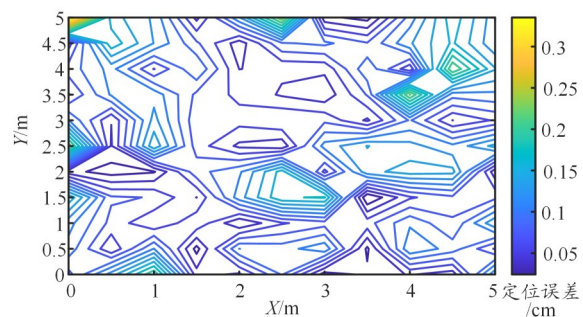
表3 不同模型消融实验定位性能对比

模型	定位网络	最大误差/cm	平均误差/cm
WGAN-GP	MobileNet V2	26.31	10.14
AT-WGAN-GP		24.93	7.97
WGAN-GP	ShuffleNet V2	35.15	8.99
AT-WGAN-GP		17.86	6.43

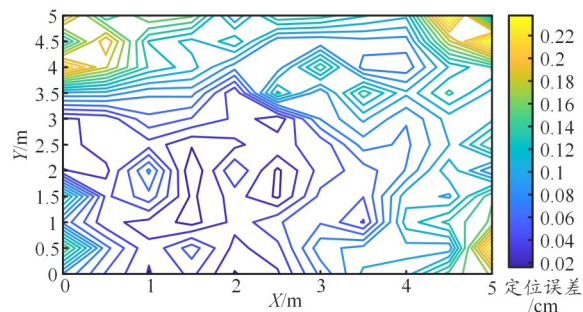
不同模型的定位误差等高线分布图如图12所示。可以看出,定位误差较大的区域主要分布在定位场景边缘。这是由于场景边缘距离灯源较远,测得的RSS值较低,从而导致定位精度下降。后续可以通过增加场景边缘原始数据或者优化自注意力机制中边缘位置的权重参数,使边缘位置的定位误差降低。



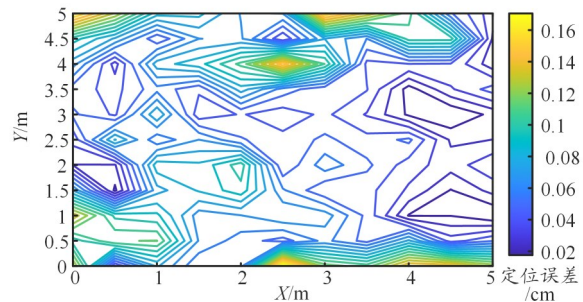
(a) WGAN-GP MobileNet V2



(b) WGAN-GP ShuffleNet V2



(c) AT-WGAN-GP MobileNet V2



(d) AT-WGAN-GP ShuffleNet V2

图12 不同模型定位误差等高线分布图

4 结束语

本文提出了AT-WGAN-GP方法,以解决指纹建库成本高昂及传统数据增强算法全局感知能力不足的问题。该方法的核心贡献在于:第一,以GP项替代权重裁剪,有效提升了WGAN-GP的训练稳定性,从根本上缓解了模式坍塌;第二,在生成器中采用渐进式降维结构并嵌入多头自注意力模块,显著增强了模型对指纹空间分布特征的全局建模能力,从而生成结构一致性更高的指纹样本;第三,在判别器中同样引入注意力机制并结合谱归一化,使其能更精准地捕捉全局细微差异,提升了判别的鲁棒性。仿真实验表明:所提AT-WGAN-GP模型在收敛速度、生成能力与数据质量上均显著优于传统GAN。在5 m×5 m×3 m的室内场景中,结合轻量化定位网络(MobileNet V2/ShuffleNet V2),即可实现平均6~8 cm的定位精度。这验证了该方法能够在保证高精度定位的同时,大幅降低指纹库构建的人力与时间成本。综上所述,本文通过融合注意力机制与WGAN-GP框架,成功实现了可见光指纹数据的高质量生成与高效位置估计,为低成本、高精度的室内VLP提供了一种有效的解决方案,具有明确的研究价值与良好的应用前景。

参考文献:

- [1] Zhu Z, Yang Y, Chen M, et al. A survey on indoor visible light positioning systems: fundamentals, applications, and challenges[J]. IEEE Communications Surveys & Tutorials, 2025, 27(3): 1656–1686.
- [2] Jian S U, Shijie C. Dynamic sitting posture recognition system using passive RFID tags in internet of things[J]. Chinese Journal of Electronics, 2025, 34(2): 586–599.
- [3] Yu M, Wan J, Liu Y, et al. Improved Bluetooth-based indoor localization for devices heterogeneity using back propagation neural network[J]. IEEE Sensors Journal, 2024, 24(17): 27763–27776.
- [4] Kong Q. Wifi-aided ultra wideband localization in indoor nlos environment[J]. IEEE Communications Letters, 2024, 28(3): 537–541.
- [5] Pagliari E, Davoli L, Ferrari G. Wi-Fi-based real-time UAV localization: a comparative analysis between RSSI-based and FTM-based approaches[J]. IEEE Transactions on Aerospace and Electronic Systems, 2024, 60(6): 8757–8778.
- [6] Kim S, Moon J, Wu J, et al. Vision-aided positioning and beam focusing for 6G terahertz communications[J]. IEEE Journal on Selected Areas

in Communications, 2024, 42(9): 2503–2519.

- [7] 王旭东,董文杰,吴楠. 基于TDOA/DOA混合的高精度室内可见光定位算法[J]. 系统工程与电子技术, 2019, 41(10): 2371–2377.
- [8] Tian J, Cong L, Qin H. Simultaneous localization and rough indoor floorplan mapping combining PDR and FM/Wi-Fi radio signals[J]. IEEE Internet of Things Journal, 2025, 12(13): 24750–24763.
- [9] 王福展,朱晓荣,陈美娟,等. 基于生成对抗网络的高精度室内无线定位方法[J]. 物联网学报, 2021, 5(2): 107–115.
- [10] 吴雅琴,陈林,侯云峰. 基于CNN-GAN的信道状态信息室内定位算法[J]. 电子测量技术, 2023, 46(24): 119–126.
- [11] 刘伟,王智豪,李卓,等. 基于cGAN-SAE的室内定位指纹生成方法[J]. 电子测量技术, 2024, 47(14): 57–63.
- [12] 曲佳,王旭东,吴楠. 基于改进ShuffleNet V2的室内可见光指纹定位算法[J]. 中国激光, 2024, 51(13): 199–211.
- [13] 赵楚韩,张洪明,宋健. 基于指纹的室内可见光定位方法[J]. 中国激光, 2018, 45(8): 202–208.
- [14] 许浩,王旭东,吴楠. 基于卷积神经网络的室内可见光指纹定位方法[J]. 激光与光电子学进展, 2021, 58(17): 227–234.
- [15] Luo Z, Li J, Zhu Y. A deep feature fusion network based on multiple attention mechanisms for joint iris-periocular biometric recognition[J]. IEEE Signal Processing Letters, 2021, 28: 1060–1064.
- [16] Joardar B K, Deshwal A, Doppa J R, et al. High-throughput training of deep CNNs on ReRAM-based heterogeneous architectures via optimized normalization layers[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2021, 41(5): 1537–1549.
- [17] Cheng Z, Yan C, Wu F X, et al. Drug-target interaction prediction using multi-head self-attention and graph attention network[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2022, 19(4): 2208–2218.
- [18] Raza A, Hashmi M F. Multiclass tumor segmentation from brain MRIs using GARU-Net: Gelu activated attention aware res-3DUNET for adaptive feature pooling[J]. IEEE Sensors Letters, 2024, 8(4): 1–4.
- [19] Samain O, Geiger B, Roujean J L. Spectral normalization and fusion of optical sensors for the retrieval of BRDF and albedo: application to vegetation, modis, and MERIS data sets[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(11): 3166–3179.
- [20] 王宁,郭梓昱,田淑珂,等. 基于融合特征t-SNE降维的控制图质量异常模式识别[J]. 系统工程理论与实践, 2024, 44(7): 2381–2393.
- [21] Pan Z, Yu W, Wang B, et al. Loss functions of generative adversarial networks (GANs): opportunities and challenges[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2020, 4(4): 500–522.
- [22] Lee I, Lee W. UniQGAN: Unified generative adversarial networks for augmented modulation classification[J]. IEEE Communications Letters, 2022, 26(2): 355–358.