

多尺度时空特征建模与自适应融合的鲁棒地磁室内定位

张志伟, 王庆虎, 刘金宇, 裴志利

(内蒙古民族大学 计算机科学与技术学院, 内蒙古 通辽 028043)

摘要: 针对地磁室内定位中特征位置语义表达能力弱、多尺度序列模糊等问题, 提出一种基于多尺度时空特征建模与自适应融合的鲁棒地磁室内定位模型, 通过挖掘地磁序列在空间域和时序域的多尺度显著特征来实现精准定位。首先, 利用特征增强将二维地磁序列重构为多维度语义表达的三维特征图, 并通过根据分治思维设计的CNN精细建模多尺度空间上下文特征; 其次, 通过显式时序语义建模, 采用局部层级时序与全局时序的建模策略, 关注序列在不同时序尺度下的动态变化, 学习其高阶长短期时序依赖关系; 最后, 采用基于注意力的自适应特征融合网络, 通过时空联合建模降低异质时空表示的位置语义冲突, 实现精准位置预测。实验结果表明, 在尺度动态变化的走廊和教室场景下, 所提方法的定位误差相较于多尺度Transformer与注意力机制结合的定位方法分别降低31.6%、29.5%, 相较于CNN与LSTM串行融合的时空定位方法分别降低37.2%、40.2%, 能够实现精准的室内定位。

关键词: 室内定位; 多尺度序列; 时空特征建模; 位置语义消歧

DOI: 10.11907/rjdk.251731

扫描二维码阅读全文:



中图分类号: TP183

文献标识码: A

文章编号: 1672-7800(2026)002-0074-09

Robust Magnetic Indoor Localization via Multi-scale Spatiotemporal Feature Modeling and Adaptive Fusion

ZHANG Zhiwei, WANG Qinghu, LIU Jinyu, PEI Zhili

(College of Computer Science and Technology, Inner Mongolia Minzu University, Tongliao 028043, China)

Abstract: Aiming at the problems of weak semantic expression ability of feature location and fuzzy multi-scale sequence in geomagnetic indoor location, a robust geomagnetic indoor location model based on multi-scale spatiotemporal feature modeling and adaptive fusion is proposed, and precise location is achieved by mining multi-scale salient features of geomagnetic sequence in spatial domain and temporal domain. Firstly, 2D geomagnetic sequence is reconstructed into 3D feature map with multi-dimensional semantic expression by feature enhancement, and multi-scale spatial context features are precisely modeled by CNN designed according to divide and conquer thinking; Secondly, through explicit time series semantic modeling, the modeling strategy of local level time series and global time series is adopted to pay attention to the dynamic changes of sequences under different time series scales, and learn their high-order long and short-term time series dependencies; Finally, the adaptive feature fusion network based on attention is used to reduce the location semantic conflict of heterogeneous spatiotemporal representations through spatiotemporal joint modeling, and achieve accurate location estimation. The experimental results show that the positioning error of the proposed method is 31.6% and 29.5% lower than that of the multi-scale transformer and attention mechanism combined positioning method, and 37.2% and 40.2% lower than that of the time-space positioning method of serial fusion of CNN and LSTM, respectively, in the corridor and classroom scenarios with dynamic scale changes, the method can achieve accurate indoor positioning.

Key Words: indoor localization; multi-scale sequence; spatiotemporal feature modeling; positional semantic disambiguation

收稿日期: 2025-10-30

基金项目: 内蒙古自然科学基金项目(2022MS06029)

作者简介: 张志伟(2001-), 男, 内蒙古民族大学计算机科学与技术学院硕士研究生, 研究方向为位置感知、普适计算; 王庆虎(1983-), 男, 硕士, 内蒙古民族大学计算机科学与技术学院副教授, 研究方向为人工智能、普适计算; 裴志利(1968-), 男, 博士, 内蒙古民族大学计算机科学与技术学院教授, 研究方向为人工智能理论及应用。本文通讯作者: 裴志利。

0 引言

基于位置的服务(Location-Based Service, LBS)在医院、商场等室内场景中的应用需求广泛^[1-2]。相较于Wi-Fi、视觉等信源,室内地磁信号具有泛在、跨时间稳定、局部可分性强等优良特性,为定位提供了可靠依据,已成为当前室内定位的重要研究方向^[3-6]。地磁序列能够同时承载空间维度的位置关联信息与时间维度的动态变化信息,这一特性极大增强了地磁序列指纹在不同位置的辨识度^[7-8]。然而,在行人移动过程中,差异化的行走速度会导致严重的多尺度序列模糊问题。如图1所示,同一路径不同速度下的采样序列在密度与特征分布上存在显著差异,降低了地磁序列的可辨识度。为解决该问题,文献[9]通过预设多组不同尺度来分割地磁序列,使用门控循环单元(Gate Recurrent Unit, GRU)提取序列中各个尺度的异常特征,并借助注意力机制进行融合,以适配不同速度下的序列变化。文献[10]采用Transformer提取地磁序列的多尺度定位特征,解决设备和行走速度差异带来的序列尺度变化问题。上述方法虽然在一定程度上缓解了多尺度序列模糊等问题,但地磁序列的划分完全依赖预设尺度,未考虑原始序列的时序连续性,容易丢失分割处的关键特征,无法保证序列特征的完整性,导致定位精度受限。

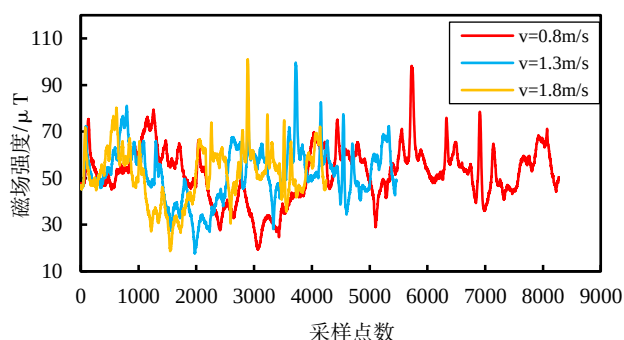


Fig. 1 Magnetic sequences sampled at different speeds

图1 不同速度下采样的地磁序列

地磁序列中隐含丰富的时空定位信息^[11]。如何挖掘深层次的空间和时序特征,降低异质时空特征的位置语义冲突,是实现室内高精度定位的关键。针对这一问题,在空间特征建模方面,文献[12]利用卷积神经网络(Convolutional Neural Network, CNN)学习磁场位置特征的空间表示,并设计投票机制融合多个CNN预测结果以提高定位精度。文献[13]通过递归图(Recurrence Plots, RPs)将地磁序列转换为二维图像特征,并使用3个递归图作为输入,结合CNN捕捉不同位置磁场的差异性空间特征,解决因相同磁场值分布导致的精度下降问题。在时序特征建模方面,文献[14]提出一种基于分层的长短期记忆(Long Short Term Memory, LSTM)网络,通过提取地磁序列的时序变化

特征进行位置分类预测。文献[15]提出基于注意力机制的LSTM定位方法,利用注意力机制为LSTM模型赋予动态权重,学习与位置相关的显著特征。在时空联合建模方面,文献[16]采用基础的CNN与LSTM串行框架,缺乏对时空特征的并行融合表示,对复杂场景下时空定位特征的学习能力不足。同时,该研究忽视了地磁序列在不同尺度下的动态表示差异,影响了定位精度。文献[17]提出使用多段连续地磁序列的定位方法,通过改进Transformer模型和联合位置估计方法,融合多个序列的时空定位特征进行定位,但序列解码阶段的误差累积限制了其预测精度。上述研究仅从单一空间域和时序域学习地磁序列定位特征,或采用空间到时序的串行融合路径,均未构建多尺度时空特征的协同建模机制,并缓解异质时空表示的位置语义冲突,难以全面挖掘地磁序列中蕴含的时空协同定位信息,在磁场分布复杂的场景中,易因特征表达能力不足而导致定位精度下降。

为解决多尺度序列模糊、时空特征建模不全面等问题,本文提出一种基于多尺度时空特征建模与自适应融合的定位网络(Multi-Scale Spatiotemporal Network, MST-Net),从空间域与时序域双视角切入,构建多尺度时空特征建模机制,并结合注意力实现时空特征自适应融合,全面解析地磁序列与位置之间的映射关系。具体而言,本文的主要创新工作与贡献点为:①针对行人移动速度差异导致的地磁序列多尺度模糊问题,提出一种多尺度空间特征建模方法。首先,通过特征增强将二维地磁序列重构为具备多维度语义表达能力的三维特征图,克服地磁序列在特征表达维度单一、语义信息不足方面的局限性;其次,使用基于分治思维设计的CNN对三维特征图的不同维度建模,学习地磁序列在不同尺度下的空间表示;最后,引入跨尺度注意力,通过动态分配注意力权重以强化关键尺度特征,抑制冗余信息,实现多尺度、多维度特征的交互融合,以构建地磁序列的多尺度空间上下文特征表示;②为全面挖掘地磁序列中隐含的深层次时空定位信息,解决地磁特征语义表达能力弱的问题,本文进一步从时序域视角出发,提出一种多尺度时序特征建模方法。首先,通过显式时序语义建模将子序列划分为多个非重叠的子段序列,避免信息冗余和序列连续特征丢失;其次,采用局部层级时序和全局时序建模方式学习每个子段之间的局部关联信息和子序列的全局信息,关注序列在不同尺度下的动态变化,捕捉序列的多尺度时序特征;最后,通过基于注意力的自适应特征融合网络实现序列空间与时序表示的联合建模,降低异质时空表示的位置语义冲突,实现精准位置预测;③在高校教学楼中进行测试实验,系统评估MST-Net模型在不同场景和移动速度下的定位效果。实验结果表明,在尺度动态变化的场景下,所提方法相比其他室内定位方法,定位精度具有明显优势。

1 系统框架

本文方法的定位流程如图2所示,主要包括离线训练与在线定位两个阶段。在离线训练阶段,使用动态连续采集方式,水平方向手持移动设备沿实验路径以高频率采样、恒定低速行走,用于采集高密度地磁序列数据。受文献[18]启发,为降低数据采集成本并模拟不同速度下地磁序列的尺度变化,对高密度地磁序列进行数据增强,构建含多种速度变化的训练数据。此外,采用滑动窗口机制将地磁序列 M 划分为连续地磁子序列。设定滑动窗口尺寸为 η , 滑动步长为 τ_1 , 从标准化处理后的地磁序列 M 中依次生成具有特定重叠部分的子序列样本 $s_i \in \mathbb{R}^{\eta \times H}$, 其中 H 表示地磁观测信号的3个正交轴向分量, 子序列 s_i 对应的标签为该序列终点处的二维坐标位置 $P_i = (x_i, y_i)$ 。在在线定位阶段, 实验人员以相同姿态采集地磁序列数据, 经过与训练阶段一致的标准化和滑动窗口处理后, 将得到的子序列样本输入已训练完成的定位模型, 预测设备当前所在位置的二维坐标。

2 MST-Net 定位模型

MST-Net 模型综合考虑地磁序列在空间域和时序域的多尺度特征表达, 通过多尺度时空特征建模与自适应特征融合来建立地磁序列与位置坐标之间的映射关联。以下介绍 MST-Net 模型中各子网络的结构设计, 系统阐述多

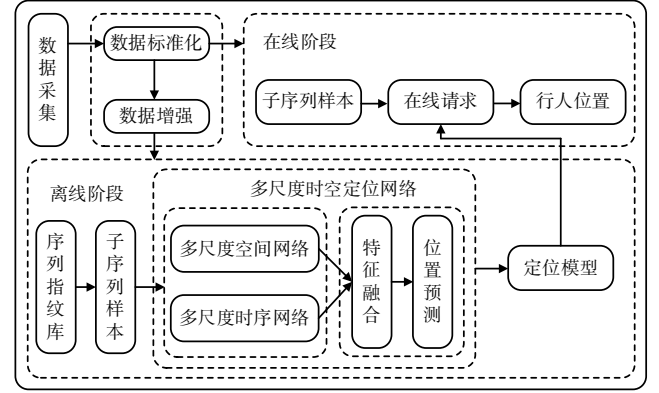


Fig. 2 Localization process

图2 定位流程

尺度地磁序列在时空双视角下的特征建模机制与融合表达。

2.1 多尺度空间特征建模

针对地磁序列存在的多尺度模糊问题, 本文提出一种多尺度空间特征建模方法。如图3所示, 该方法主要由3个部分组成: ①将二维地磁序列重构为多维度语义表达的三维特征图, 显式保留地磁序列在空间、轴向与时序维度的多维耦合特性, 克服地磁序列在特征刻画能力方面的局限性; ②设计基于分治思想的CNN, 采用定向卷积操作与卷积核分解设计, 通过对不同维度的特性进行针对性建模来逐级提取地磁序列在不同尺度和维度下的语义信息; ③通过跨尺度特征融合机制实现多尺度、多维度特征表示的深度融合, 为后续位置预测提供辨识度更高的多尺度空间特征基础。

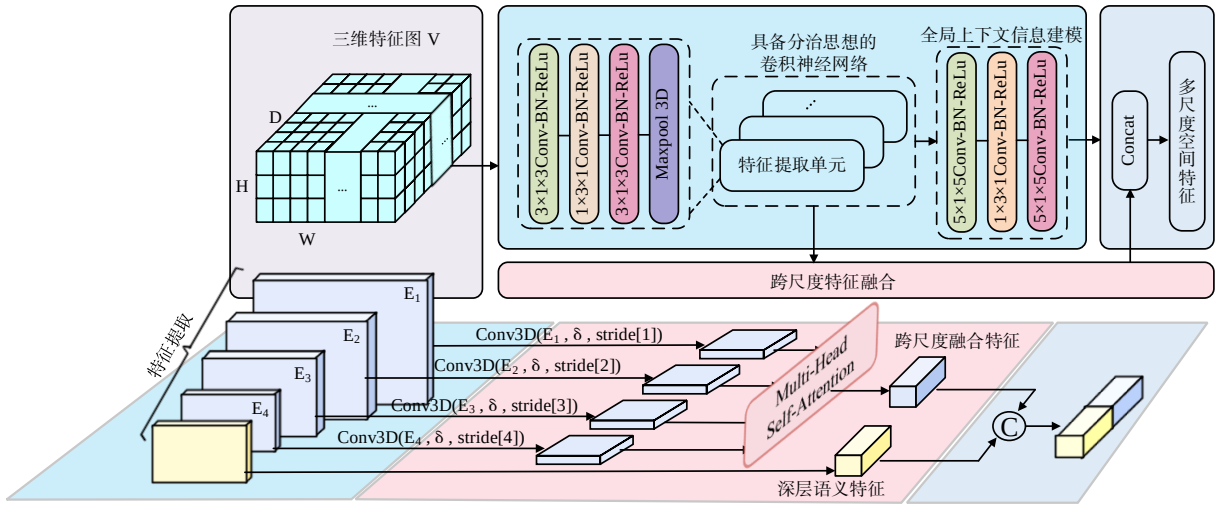


Fig. 3 Multi-scale spatial feature modeling

图3 多尺度空间特征建模

2.1.1 地磁序列三维特征表示建模

针对现有地磁序列在特征表达能力上的局限性, 尤其是在建模非连续位置之间地磁信号观测值的关联信息方面表现出的不足, 本文提出一种地磁序列三维重构方法, 旨在充分挖掘序列中蕴含的空间特征、时序特征以及多轴交互关系, 为后续空间建模任务提供更具表达力的特征基

础。具体来说, 对于给定地磁子序列 s_i , 使用滑动窗口对其进行分片处理, 设定滑动窗口尺寸为 ξ , 滑动步长为 τ_2 , 可从原始地磁子序列中截取 D 个连续片段。表示为:

$$D = (\eta - \xi) / \tau_2 + 1 \quad (1)$$

式中: 每个子片段的维度为 $\mathbb{R}^{H \times \xi}$ 。对所有片段沿深度方向进行堆叠, 构建出维度为 $D \times H \times W$ 的三维特征图 V ,

其中空间维度 D 表示局部观测片段的堆叠顺序,反映出采样路径上的相对位置结构;轴向维度 H 表示三轴地磁分量;时序维度 W 表示每个片段内部的地磁信号动态变化趋势,体现了局部时序演化规律。

2.1.2 基于分治思想的 CNN

普通 CNN 在同时建模空间—轴向—时序依赖关系时,存在对各维度特征捕捉不充分、针对性不足的问题。为此,设计一种基于分治思想的 CNN,采用定向卷积操作与卷积核分解设计,沿不同维度对三维卷积核进行拆解,通过分而治之的策略分别精准建模三维特征图 V 的局部空间依赖、片段内时序演化以及地磁三轴分量的交互关系,提升网络对各维度信息的学习能力。具体来说,该网络包含 γ 层特征提取单元 Φ , 每层 Φ 均由以下 3 个卷积子层构成:

(1) 在空间—时序建模子层利用时空卷积核分别作用于空间维 D 和时序维 W , 轴向维 H 则独立处理。该层旨在对三维特征图 V 的空间维,以稀疏采样模式精准捕获其非连续位置之间地磁信号观测值的远程关联信息。对时序维度执行连续卷积操作,细致建模每个片段内部的时序动态演化规律。轴向维度则对地磁三轴分量独立处理,避免特征在该阶段被过度加工,保留其固有信息,为后续跨轴交互子层的轴间关系学习提供高质量特征输入。表示为:

$$T_{D,W} = \text{ReLU}(\text{BN}(\text{Conv3D}(V, \varphi_{D,W}))) \quad (2)$$

式中: $\varphi_{D,W}$ 表示作用于空间维 D 和时序维 W 的卷积核; Conv3D 表示三维卷积运算; BN 表示批量归一化; ReLU 表示激活函数。

(2) 为学习地磁三轴分量之间存在的非线性耦合特性,本文构建跨轴交互建模子层,对轴向维度进行特征交互建模,基于上一建模子层输出的浅层时空特征 $T_{D,W}$, 引入专用于学习地磁三轴分量间交互关系的轴向卷积核,实现跨轴特征融合。表示为:

$$T_H = \text{ReLU}(\text{BN}(\text{Conv3D}(T_{D,W}, \varphi_H))) \quad (3)$$

式中: φ_H 表示作用于轴向维度的卷积核。通过学习地磁三轴分量之间的相互作用关系,模型对地磁环境中复杂磁场变化的建模能力可得到有效提升。

(3) 为进一步精炼时空相关特征,基于上述两个卷积子层的建模,本文构建强化时空融合子层,通过再次引入时空卷积核聚焦空间与时序维度,在浅层时空与轴间关系交互特征的基础上挖掘深层次时空信息,实现时空特征与轴间交互关系的同步学习。此外,为增强模型对全局语义信息的感知能力,针对特征提取单元 Φ 输出的高阶特征 E_4 , 本文设计全局上下文信息建模单元 Ψ 。通过引入一组大尺寸卷积核 $v[i], i \in [1, 3]$ 显著扩大感受野,充分学习跨片段的长期依赖关系,并持续优化三轴分量交互建模,得到富含全局上下文信息的深层次语义特征 Θ 。表示为:

$$\Theta = \Psi(E_4, v[i]) \quad (4)$$

2.1.3 跨尺度特征融合

考虑到不同层级间的特征分别承载着不同的尺度信息,本文设计了基于跨尺度注意力的特征融合机制,对不同层级之间的特征进行融合,挖掘尺度间的依赖关系,并根据动态权重分配突出关键尺度特征,抑制冗余信息。首先,针对特征提取单元 Φ 输出的不同尺度特征集合 $\varpi = \{E_1, E_2, E_3, E_4\}$ 分别设计 4 组跨步卷积,将其映射为相同空间尺寸,形成统一融合空间 Ω 。表示为:

$$\Omega = \text{Conv3D}(\varpi, \delta, \text{stride}[i]), i \in [1, 4] \quad (5)$$

式中: δ 为卷积核参数; stride 为卷积步幅约束。

其次,将 Ω 中经过对齐的特征沿通道维度按顺序拼接,形成信息完整的特征序列 ρ 。利用跨尺度注意力学习不同尺度特征之间的依赖关系,以挖掘不同层级特征间的潜在关联。具体来说,通过线性变换将特征序列 ρ 映射为对应的查询向量 q 、键向量 k 、值向量 v 。公式为:

$$q, k, v = \rho W_q, \rho W_k, \rho W_v \quad (6)$$

式中: W_q, W_k, W_v 为可学习参数矩阵。

通过查询向量 q 与键向量 k 的转置进行点乘运算,计算尺度间的相关性得分矩阵。公式为:

$$A = \frac{qk^T}{\sqrt{d_k}} \quad (7)$$

式中: 矩阵 A 中的每个元素 $a_{i,j}$ 均代表某一尺度特征的查询向量与另一尺度特征键向量之间的相似性。对相关性得分进行 Softmax 归一化。公式为:

$$\alpha_{i,j} = \frac{e^{a_{i,j}}}{\sum_{t=1}^N e^{a_{i,t}}} \quad (8)$$

式中: α 为注意力权重矩阵,且权重之和为 1,权重值越大,表示对应尺度特征对当前融合结果的贡献度越高。基于 α 对值向量 v 进行加权求和,得到尺度融合特征。表示为:

$$\Lambda_i = \sum_{j=1}^N \alpha_{i,j} v_j \quad (9)$$

式中: Λ_i 表示融合了不同尺度依赖关系的特征,体现出跨尺度特征之间的动态交互与选择性聚合。

最后,将跨尺度融合特征 Λ 与深层次语义特征 Θ 沿通道拼接,形成具备全局感知能力与跨尺度依赖关系的统一空间语义表示,改善地磁序列在特征提取中的表达局限,为后续位置预测任务提供信息更加丰富的特征基础。

2.2 多尺度时序特征建模

为挖掘地磁序列中隐含的深层次时空定位信息,进一步从时序域视角出发,提出一种基于双分支 Transformer 结构的多尺度时序特征建模方法。如图 4 所示,首先通过显式时序语义建模将子序列划分为多个非重叠子段序列,避免信息冗余和序列连续特征丢失;其次利用双分支 Transformer 网络分别对子段序列和子序列进行局部层级时序和全局时序建模,学习跨子段序列间的特征关联信息与全局时序特征,全面提取地磁序列中蕴含的多尺度时序特征。

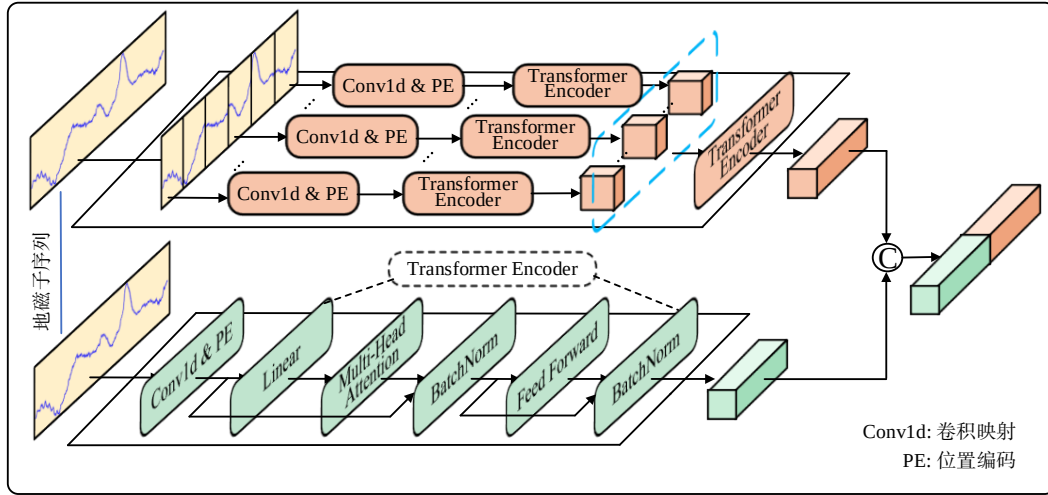


Fig. 4 Multi-scale temporal feature modeling

图4 多尺度时序特征建模

2.2.1 地磁序列分段建模

行人运动中采集的地磁序列数据受速度、高度随机性影响,呈现出显著的多尺度特性。Transformer模型的自注意力机制对序列局部上下文信息的敏感性较弱,若直接对地磁子序列进行学习,易导致局部关键信息丢失和全局序列特征误匹配^[19]。为此,本文对地磁序列进行非重叠均衡分段,通过显式时序语义建模,在重叠子序列的基础上得到非重叠子段序列,既能规避子段间的信息冗余,避免序列连续特征丢失,又能增强模型对序列局部时序变化的学习能力,为后续多尺度时序特征提取奠定基础。具体来说,对于给定子序列 s_i ,将其均匀划分为 μ 个长度为 $\lambda = \eta/\mu$ 的非重叠子段序列。定义为:

$$Q = \{q_j | q_j = x_{\lambda(j-1): \lambda j}, j \in [1, \mu]\} \in s_i \quad (10)$$

式中: Q 表示子序列 s_i 分段后的子段序列 q_j 的集合, $q_j \in \mathbb{R}^{\lambda \times H}$; x 表示地磁序列中单个时间步的地磁数据。考虑到分段操作可能在子段边界处造成时序信息损失,本文通过引入子序列全局建模分支来提供完整的子序列上下文语义补全,用于弥补边界信息缺失,实现对子序列多尺度时序特征的充分挖掘。

2.2.2 双分支Transformer网络

基于局部子段序列集合 $Q = \{q_1, q_2, \dots, q_\mu\}$ 与子序列 s_i 构建局部层级时序分支网络与全局时序分支网络,在实现子段间语义信息聚合的基础上融合子序列的全局特征,以实现多尺度时序特征的融合学习。需要说明的是,室内定位任务是从输入序列中提取表征位置的定位特征,进而预测位置,属于序列到单点的回归任务;而标准的Transformer采用编码器与解码器结构,主要面向序列到序列的生成任务^[20]。因此,本文提出的双分支Transformer网络仅保留编码器结构,在更契合任务的同时降低模型复杂度、缓解过拟合风险,并提升训练与推理效率^[21]。

首先,以序列长度为 w 的子段 $q_j \in \mathbb{R}^{w \times 3}$ 作为局部层级

时序分支网络的输入,构成包含 w 个特征向量 x_i 的序列。表示为:

$$X \in \mathbb{R}^{w \times 3} = [x_1, x_2, \dots, x_i, \dots, x_w] \quad (11)$$

$$\text{式中: } x_i \in \mathbb{R}^3 = [m_x, m_y, m_z].$$

通过一维卷积将特征向量 x_i 映射至 d 维向量空间,得到隐层表示 u_i 。投影后的序列为:

$$U \in \mathbb{R}^{w \times d} = [u_1, u_2, \dots, u_i, \dots, u_w] \quad (12)$$

为保留序列的时序信息,引入正余弦位置编码 $W_{pos} \in \mathbb{R}^{w \times d}$,与投影结果相加得到输入序列的增强表示 U' 。表示为:

$$U' = U + W_{pos} = [u'_1, u'_2, \dots, u'_i, \dots, u'_w] \quad (13)$$

将 U' 输入一级段内Transformer编码器,用于提取子段内部的局部时序依赖关系。该编码器由自注意力子层和前馈神经网络子层组成,采用残差连接与层归一化操作缓解梯度消失问题并促进特征累积,稳定提取子段内部的局部时序特征 Γ_θ 。随后,二级段间Transformer编码器对所有 Γ_θ 再次建模,学习子段间的时序关联信息,实现跨子段语义聚合与尺度间信息交互,形成段间语义表示 Γ_ϕ 。

其次,全局时序分支网络旨在弥补局部层级时序建模中因划分子段 q_j 而产生的边界时序信息损失。该分支以子序列 s_i 作为输入,先通过一维卷积投影将序列特征映射至适配维度,再融入位置编码保留时序信息,随后送入Transformer编码器中,学习隐含的长期时序依赖关系,提取全局时序特征 Γ_g 。

最后,将局部层级时序网络输出的段间语义表示 Γ_ϕ 与全局时序网络提取的子序列全局时序特征表示 Γ_g 拼接融合,形成序列的多尺度时序特征表示。该特征同时整合序列的局部短时与全局长时依赖关系,为后续位置预测任务提供可靠的多尺度时序定位特征。

2.3 自适应特征融合与位置预测

针对地磁序列的时空特征融合问题,本文采用基于注

意力的自适应特征融合网络,通过动态加权实现空间与时序特征的互补融合,形成兼具局部细节与全局关联的序列多尺度时空表示。具体而言,设多尺度空间特征为 F_s ,多尺度时序特征为 F_t ,将其拼接得到时空联合特征表示:

$$F_j = [F_s; F_t] \quad (14)$$

式中: F_j 涵盖地磁序列的空间位置分布与时序演化信息,为注意力权重计算提供完整的特征基础。其经线性变换与 $ReLU$ 激活函数处理后输入至 $Softmax$ 函数,获得空间与时序特征的归一化注意力权重向量。表示为:

$$\alpha = Softmax(ReLU(W \cdot F_j + b)), \alpha \in [\alpha_s, \alpha_t] \quad (15)$$

式中: α_s, α_t 分别表示空间和时序特征的权重。

最终多尺度时空融合特征可表示为:

$$F_f = \alpha_s \cdot F_s + \alpha_t \cdot F_t \quad (16)$$

该机制可自适应调整两类特征的贡献度,突出判别性信息并抑制冗余消息,实现空间与时序特征的互补融合,降低异质时空表示的位置语义冲突。

将多尺度时空联合特征 F_f 展平后输入由多层感知机构成的位置预测网络,解算最终的二维坐标 $(\hat{x}, \hat{y})$ 。模型训练过程中采用均方误差 (Mean Square Error, MSE) 作为损失函数,衡量预测坐标与真实坐标的偏差。计算公式为:

$$MSE(x, y, \hat{x}, \hat{y}) = \frac{1}{N} \sum_{i=1}^N ((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2) \quad (17)$$

式中: (x_i, y_i) 为第 i 个样本的真实坐标; $(\hat{x}_i, \hat{y}_i)$ 为对应的预测坐标; N 为样本总数。该损失函数可有效惩罚较大的坐标预测误差,引导模型学习更精准的映射规则。

3 实验方法与结果分析

3.1 实验设置与评估标准

为系统评估 MST-Net 模型在尺度动态变化测试场景下的定位表现,在真实室内环境中设计并开展了多组实验。如图 5 所示,选取高校教学楼中相对开阔的走廊区域和相对封闭的教室区域作为实验环境,模拟真实室内行人活动场景,验证模型在不同场景下的定位效果。定位模型的基本实验参数如表 1 所示。

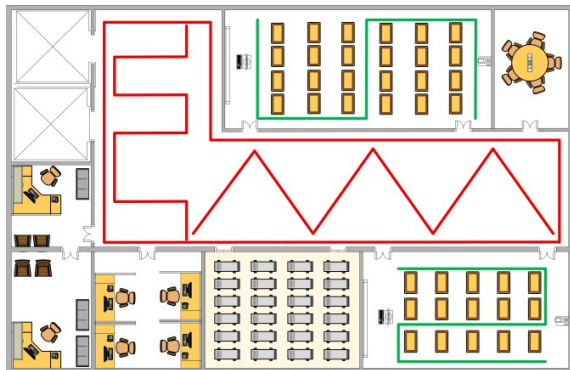


Fig. 5 Plan view of the experimental site

图 5 实验场地平面图

Table 1 Experimental parameter settings

表 1 实验参数设置

参数	值
迭代次数	300
批次大小	64
学习率	0.000 2
序列长度	400
特征提取单元层数	4

在数据采集阶段,使用 100 Hz 高采样频率、0.5 m/s 的恒定低速采集训练数据,以确保地磁数据的细粒度特征不丢失、样本分布均匀,为模型学习细粒度地磁时空定位信息提供高质量样本。在采集测试数据时,将设备的采样频率调整至 50 Hz,并参照文献[22]中的行人运动速度划分模式,覆盖 0.8 ~ 1.8 m/s 的速度区间。该设置旨在模拟真实行人非匀速运动状态,客观验证定位模型在尺度动态变化场景下的适应性。

为评估模型的定位表现,采用基于欧式距离的平均定位误差 e 作为核心评估指标。该指标通过计算所有测试样本的预测位置与真实位置间欧氏距离的均值来直观量化定位误差。其值越小,表示模型的定位精度越高。计算公式为:

$$e = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \quad (18)$$

为规避实验结果的偶然性,对每条实验路径均重复测试 5 次,在每次测试时保持设备采样频率、手持姿态一致,取 5 次测试定位误差的平均值作为该实验路径的最终定位误差值。

3.2 比较实验

为验证 MST-Net 模型的有效性,在相同实验场景与速度条件下将其与两种先进室内定位方法进行比较,分别为:①MS-Trans^[10]。其为基于多尺度 Transformer 与注意力机制的定位方法,通过预设多组尺度来分割地磁序列,并对每个子序列赋予相应的权重,以提取地磁序列的多尺度显著特征,是解决行人移动速度差异导致的多尺度序列模糊问题的先进方法;②CNN-LSTM^[16]。其为采用串行结构的时空融合定位方法,通过 CNN 与 LSTM 分别提取地磁序列的时空特征以实现准确定位,是使用地磁序列进行时空特征融合的先进方法。表 2 为 3 种方法的平均定位误差。可以看出,无论是走廊还是教室场景的匀速、变速条件下, MST-Net 模型的平均定位误差均低于 MS-Trans 和 CNN-LSTM 模型。

Table 2 Average localization errors of different localization methods

表 2 不同定位方法的平均定位误差 (m)

定位方法	走廊		教室	
	匀速	变速	匀速	变速
MST-Net	0.52	0.54	0.51	0.55
MS-Trans ^[10]	0.73	0.79	0.71	0.78
	(↓ 28.8%)	(↓ 31.6%)	(↓ 28.2%)	(↓ 29.5%)
CNN-LSTM ^[16]	0.80	0.86	0.82	0.92
	(↓ 35.0%)	(↓ 37.2%)	(↓ 37.8%)	(↓ 40.2%)

图6为3种方法的平均定位误差累积分布函数(Cumulation Distribution Function, CDF)。可以看出,在两种测试场景的匀速和变速条件下,MST-Net模型的平均定位误差分布均呈现出明显优势。一方面,在平均定位误差小于1 m的区间内,MST-Net模型的CDF曲线累积概率更高,意味着其在高精度区间的定位结果占比更大;另一方面,与两种对照模型相比,MST-Net模型的定位误差最大值明显更低,有效避免了极端定位误差现象,进一步验证了该模型的有效性、稳定性。究其原因,MST-Net模型可深度挖掘地磁序列在空间域和时序域的多尺度显著特征,重点关注序列在不同尺度下的动态变化规律,并结合注意力机制降低异质时空表示的位置语义冲突,完成跨时空特征的动态融合。其核心优势在于对多尺度和时空域的协同建模:在空间域,模型依托具备分治思想的CNN和注意力机制,将不同尺度下的空间特征深度融合,确保空间信息的完整

性和关联性;在时序域,其将子序列划分为多个非重叠的子段序列,以避免信息冗余和序列连续特征丢失,并结合局部层级时序和全局时序的建模方式同步提取序列的局部和全局信息,得到序列的多尺度时序特征。相比之下,MS-Trans通过预设多组尺度分割地磁序列,虽然在一定程度上缓解了多尺度序列模糊的问题,但预设尺度难以适配不同场景变化,且会破坏地磁序列特征的完整性并引入噪声,导致定位精度受限。CNN-LSTM虽然通过CNN与LSTM结合的方式提取地磁序列的时空融合特征,提升了模型的定位精度,但是缺乏多尺度特征建模机制,忽视了序列在不同尺度下的动态变化,且未能实现空间与时序特征的深层次建模与特征融合,难以充分挖掘地磁序列中蕴含的丰富时空定位信息,最终制约了模型的特征提取能力。

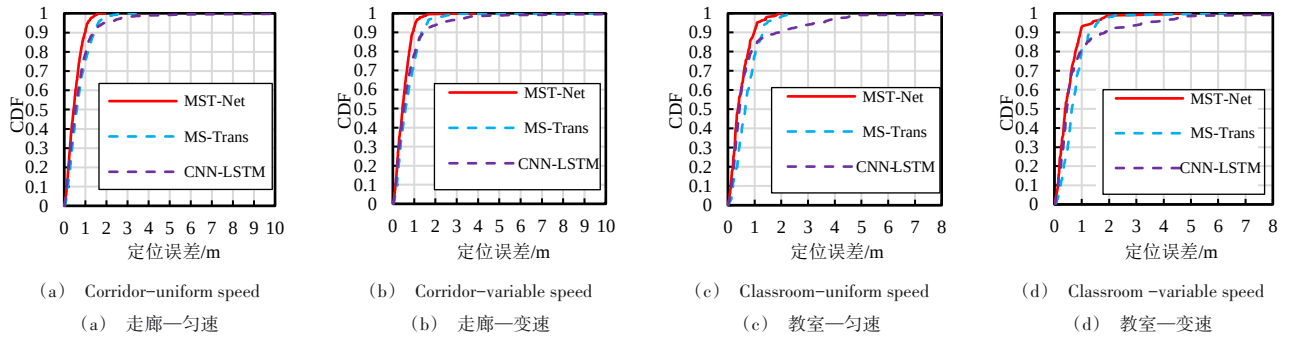


Fig. 6 CDF of average localization errors for different localization methods

图6 不同定位方法的平均定位误差 CDF

3.3 消融实验

为进一步探究 MST-Net 模型中各特征建模网络的有效性,本文设计两种变体模型,分别为仅保留多尺度空间特征建模网络的 MS-Net 与仅含有多尺度时序特征建模网络的 MT-Net。消融实验结果如图7和图8所示,可以看出,MST-Net模型在所有测试条件下均表现出最佳定位效果,其平均定位误差显著低于 MS-Net 模型和 MT-Net 模

型。由图7可知,在走廊和教室场景的匀速、变速条件下,MST-Net模型在0.5 m内的定位误差比例分别为55.8%、54.9%、60.4%、58.2%,MS-Net模型分别为35.9%、32.4%、47.2%、40.9%,MT-Net模型分别为38.7%、38.8%、47.6%、39.1%,说明MST-Net模型的定位精度相较变体模型显著提升。

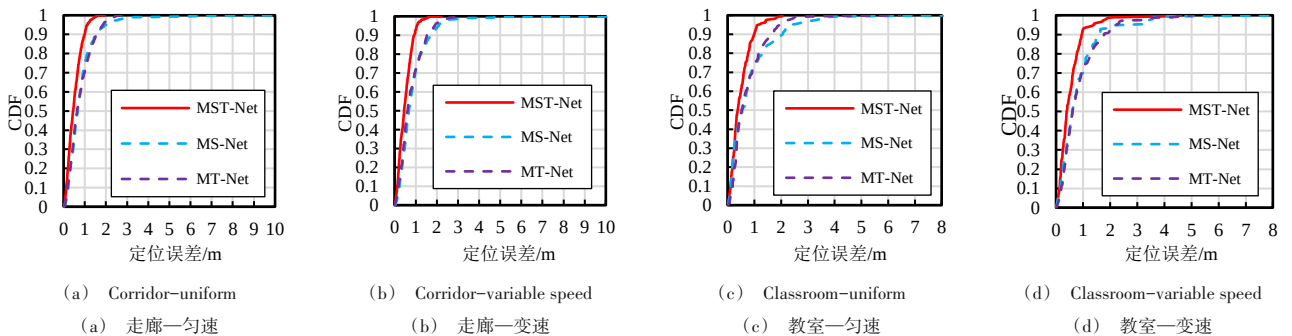


Fig. 7 CDF of average localization errors for MST-Net and its variant models

图7 MST-Net及其变体模型的平均定位误差 CDF

由图8可知,在走廊场景的匀速、变速条件下,MST-Net的平均定位误差分别为0.52 m、0.54 m,而MS-Net和

MT-Net的平均定位误差分别为0.79 m、0.86 m和0.74 m、0.77 m。相比之下,MST-Net模型的平均定位误差分别降

低 34.2%、37.2% 和 29.7%、29.9%。在教室场景的匀速、变速条件下,MST-Net 的平均定位误差分别为 0.51 m、0.55 m,而 MS-Net 与 MT-Net 的平均定位误差分别为 0.82 m、0.91 m 和 0.78 m、0.81 m。相比之下,MST-Net 模型的平均定位误差分别降低 37.8%、39.6% 和 34.6%、32.1%。实验结果表明,单一 MS-Net 或 MT-Net 对地磁序列的学习能力存在局限性。MS-Net 仅依赖空间特征,易受速度变化导致的序列采样密度波动影响,导致定位误差较大。MT-Net 仅依赖时序特征,虽然在变速情况下的误差增幅较小,但整体定位精度较低。MST-Net 则通过多尺度时空联合建模,同时考虑地磁序列在空间域和时序域的多尺度特征表达,实现了平均定位误差的显著降低。

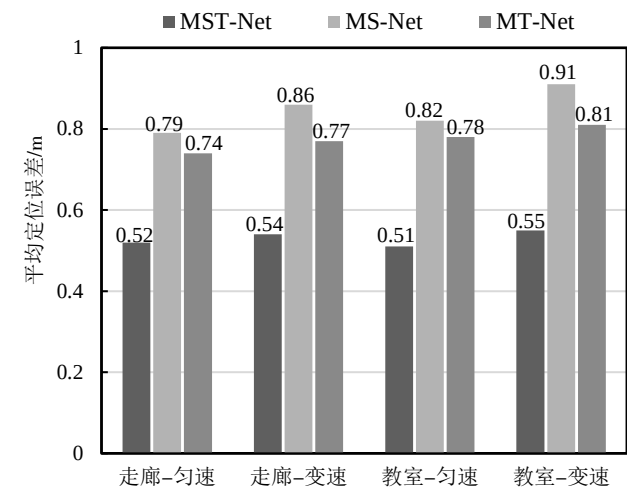


Fig. 8 Average localization errors of different variant models

图 8 不同变体模型的平均定位误差

3.4 参数实验

为研究关键超参数对模型定位精度的调控机制,首先探究序列长度(即滑动窗口的大小)这一核心参数对模型定位结果的影响。图 9 直观呈现了 MST-Net 模型在不同序列长度下的平均定位误差变化趋势。可以看出,随着序列长度的逐步增加,该模型的平均定位误差呈下降态势。这是由于较长的地磁序列蕴含更丰富的时空定位信息,可使模型更精准地学习序列与位置之间的映射关系,提升定位精度。当序列长度超过 400 时,模型平均定位误差的下降速率明显放缓,这表明此时序列所包含的地磁信息已满足模型实现高精度定位的核心需求,继续增加序列长度所带来的时空信息增益显著衰减。同时,过长的地磁序列会导致模型计算开销增加,综合权衡定位精度与效率,本文选择序列长度=400 作为模型输入。

在以上研究的基础上进一步聚焦多尺度空间特征建模网络的核心结构——特征提取单元。表 3 展示了不同层数特征提取单元对模型定位效果的影响。可以看出,随着层数的增加,模型的特征表达能力逐步增强,平均定位误差逐渐降低。当层数增加至 4 时,模型在匀速与变速条件下的平均定位误差均达到最低。然而,若继续增加特征

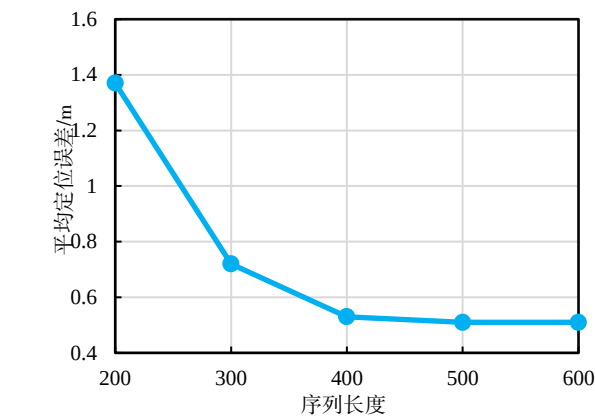


Fig. 9 Average localization errors at different sequence lengths

图 9 不同序列长度下的平均定位误差

提取单元的层数,平均定位误差反而呈现回升趋势,这表明过深的网络结构会导致模型陷入过度拟合困境,削弱其对位置坐标的预测能力。因此,综合权衡模型复杂度与定位精度,本文选取特征提取单元的层数为 4,既能确保模型充分发挥特征提取能力,又可规避过拟合风险。

Table 3 Average localization errors of different feature extraction unit layers

特征提取单元层数	平均定位误差/m		
	匀速	变速	平均
2	0.63	0.65	0.64
3	0.54	0.56	0.55
4	0.52	0.55	0.53
5	0.57	0.58	0.57
6	0.58	0.63	0.61

此外,本文还探究了不同学习率对模型定位效果的影响,结果如表 4 所示。可以看出,当学习率为 0.000 2 时,模型的匀速和变速平均定位误差分别为 0.52 m、0.55 m,方差分别为 0.17 m²、0.21 m²,相较于其他学习率时表现更佳。当学习率较大时,模型容易在损失函数的最优解附近震荡,难以稳定收敛。而当学习率过小时,模型容易陷入次优解,同样无法稳定收敛。因此,本文将学习率设置为 0.000 2,使模型在保持高精度的同时也具有一定的稳定性。

Table 4 Influence of learning rate on model localization effect

表 4 学习率对模型定位效果的影响

学习率	平均定位误差/m		方差/m ²	
	匀速	变速	匀速	变速
0.002	0.77	0.85	0.18	0.23
0.001	0.70	0.74	0.19	0.21
0.000 2	0.52	0.55	0.17	0.21
0.000 1	0.75	0.79	0.21	0.21
0.000 02	0.78	0.88	0.25	0.37

4 结语

本文针对地磁室内定位任务中存在的地磁特征语义

表达能力弱和多尺度序列模糊问题,提出一种多尺度时空特征建模与自适应融合的鲁棒室内定位模型 MST-Net,通过挖掘地磁序列在空间域与时序域的多尺度特征表达来实现高精度定位。首先,在空间域通过特征增强对二维地磁序列进行三维重构,以得到多维度语义表达的三维特征图,并通过分治思维设计的 CNN 和跨尺度特征融合机制来建模多尺度空间上下文特征;其次,从时序域视角切入,通过显式时序语义建模,并采用局部层级时序与全局时序结合的方法关注序列在不同尺度下的动态变化,捕捉其隐含的多尺度时序特征;最后,借助自适应特征融合网络实现跨时空特征的动态加权融合,降低异质时空表示的位置语义冲突,增强模型对地磁序列变化的适应能力与特征判别能力。实验结果表明,MST-Net 模型在尺度动态变化测试场景中较现有先进方法具有更高的定位精度与鲁棒性。未来将进一步对模型的场景迁移性展开研究,探索其对更多复杂室内场景的适应性以及不同场景间定位的无缝衔接技术,提升模型对不同室内场景的普适能力。

参考文献:

- [1] ZAFARI F, GKELIAS A, LEUNG K K. A survey of indoor localization systems and technologies[J]. IEEE Communications Surveys & Tutorials, 2019, 21(3): 2568–2599.
- [2] HUANG H, GARTNER G, KRISP J M, et al. Location based services: ongoing evolution and research agenda[J]. Journal of Location Based Services, 2018, 12(2): 63–93.
- [3] HAO L, HUANG B, JIA B, et al. DHCLoc: a device-heterogeneity-tolerant and channel-adaptive passive WiFi localization method based on DNN[J]. IEEE Internet of Things Journal, 2022, 9(7): 4863–4874.
- [4] TSENG P Y, LIN J J, CHAN Y C, et al. Real-time indoor localization with visual SLAM for in-building emergency response[J]. Automation in Construction, 2022, 140: 104319.
- [5] SUN M, WANG Y, XU S, et al. Indoor geomagnetic positioning using the enhanced genetic algorithm-based extreme learning machine[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1–11.
- [6] LIU Y X, LI W, WEI D Y, et al. Online calibration method of smartphone magnetometer in vehicle geomagnetic matching positioning [C]//2022 IEEE 12th International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2022: 1–8.
- [7] MA S, PEI K, QI H Y, et al. Research on geomagnetic indoor high-precision positioning algorithm based on generative model[J]. Journal on Communications, 2023, 44(6): 211–222.
马帅,裴科,祁华艳,等. 基于生成模型的地磁室内高精度定位算法研究[J]. 通信学报, 2023, 44(6): 211–222.
- [8] PU Q L, CHEN Y K, ZHOU M, et al. Indoor WiFi localization algorithm based on the improved contrastive learning and parallel fusion neural network [J]. Chinese Journal of Scientific Instrument, 2024, 45(1): 101–110.
- 蒲巧林,陈有坤,周牧,等. 基于改进对比学习和并行融合神经网络的室内 WiFi 定位算法[J]. 仪器仪表学报, 2024, 45(1): 101–110.
- [9] NIU Q, HE T, LIU N, et al. MAIL: multi-scale attention-guided indoor localization using geomagnetic sequences[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(2): 1–23.
- [10] WANG Q, WANG L, FU M, et al. Multiscale transformer and attention mechanism for magnetic spatiotemporal sequence localization[J]. IEEE Internet of Things Journal, 2024, 11(11): 19454–19469.
- [11] BAE H J, CHOI L. Large-scale indoor positioning using geomagnetic field with deep neural networks[C]//2019 IEEE International Conference on Communications (ICC), 2019: 1–6.
- [12] ASHRAF I, KANG M, HUR S, et al. MINLOC: magnetic field patterns-based indoor localization using convolutional neural networks[J]. IEEE Access, 2020, 8: 66213–66227.
- [13] ABID M, LEFEBVRE G. Improving indoor geomagnetic field fingerprinting using recurrence plot-based convolutional neural networks[J]. Journal of Location Based Services, 2021, 15(1): 61–87.
- [14] WANG L, LUO H, WANG Q, et al. A hierarchical LSTM-based indoor geomagnetic localization algorithm[J]. IEEE Sensors Journal, 2022, 22(2): 1227–1237.
- [15] YANG Y, CUI L Z, GUO Q Q, et al. LSTM geomagnetic positioning algorithm based on Attention mechanism in underground coal mine[J]. Bulletin of Surveying and Mapping, 2022(1): 72–78.
杨勇,崔丽珍,郭倩倩,等. 煤矿井下基于 Attention 机制的 LSTM 地磁定位算法[J]. 测绘通报, 2022(1): 72–78.
- [16] DING X, ZHU M, XIAO B. Accurate indoor localization using magnetic sequence fingerprints with deep learning [C]// International Conference on Algorithms and Architectures for Parallel Processing, 2022: 65–84.
- [17] ZHUANG H, HE T, NIU Q, et al. Efficient indoor localization with multiple consecutive geomagnetic sequences [C]//2022 International Conference on Computer Communications and Networks, 2022: 1–9.
- [18] WANG Q, JIA J, DENG Y, et al. DarLoc: deep learning and data-feature augmentation based robust magnetic indoor localization[J]. Expert Systems with Applications, 2024, 244: 122921.
- [19] LI S, JIN X, XUAN Y, et al. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting [DB/OL]. <https://arxiv.org/abs/1907.00235>.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [DB/OL]. <https://arxiv.org/pdf/1706.03762v6>.
- [21] ZERVEAS G, JAYARAMAN S, PATEL D, et al. A transformer-based framework for multivariate time series representation learning [C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021: 2114–2124.
- [22] MURTAGH E M, MAIR J L, AGUIAR E, et al. Outdoor walking speeds of apparently healthy adults: a systematic review and meta-analysis[J]. Sports Medicine, 2021, 51(1): 125–141.

(责任编辑:尹晨茹)