



雲南財經大學  
YUNNAN UNIVERSITY OF FINANCE AND ECONOMICS

## 本科学士生毕业论文（设计）

题 目 面向复杂环境的地面介质识别算法研究

学 院	<u>信息学院</u>
专 业	<u>计算机科学与技术</u>
班 级	<u>卓计 21-1 班</u>
学 号	<u>202105001997</u>
姓 名	<u>冯昌群</u>
指 导 教 师	<u>欧新宇</u>
职 称	<u>教 授</u>


## 云南财经大学

### 本科学生毕业论文（设计）原创性及知识产权声明

本人郑重声明：所呈交的毕业论文（设计）是本人在导师的指导下取得的成果。对本论文（设计）的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。因本毕业论文（设计）引起的法律结果完全由本人承担。

本毕业论文（设计）成果归云南财经大学所有。

特此声明

毕业论文（设计）作者签名： 

作者专业：计算机科学与技术

作者学号：202105001997

2025年 5 月 16 日

## 摘 要

复杂环境中地面介质的正确识别是移动机器人自主导航系统中的关键任务。然而，传感器数据的高噪声性和不同地面介质特征的相似性给分类带来挑战。本文提出了一种基于时频域融合与邻域平滑驱动的地面介质分类算法，创新性地采用模块化设计将特征提取与分类过程解耦，使特征提取器和分类器均可独立优化和灵活替换，大大提高了系统的适应性和可扩展性。算法首先通过具有多尺度感受野的TCN时域特征提取器并设计对比损失函数提取时域特征，结合快速傅里叶变换（FFT）获取频域特征，并利用四元数距离度量构建邻域网络进行特征平滑；然后将融合特征输入支持向量机进行分类。实验结果表明，所提方法在包含9种地面类型的数据集上达到了99.13%的分类准确率，比传统方法提高了约9个百分点。此外，通过消融实验进一步验证了所提方法各模块的有效性。本研究为移动机器人环境感知提供了新的技术路线，对机器人自主导航和智能决策具有实践价值。

**关键词：** 解耦设计 时间卷积网络 对比学习 邻域平滑 时频特征融合

### Abstract

Accurate identification of ground media in complex environments is a key task in mobile robot autonomous navigation systems. However, high noise in sensor data and similarities between different ground media features pose challenges for classification. This paper proposes a ground media classification algorithm based on time-frequency features, contrastive learning, and neighborhood smoothing strategies. The algorithm innovatively adopts a modular design that decouples the feature extraction from the classification process, allowing the feature extractor and classifier to be independently optimized and flexibly replaced, greatly improving the system's adaptability and scalability. The algorithm first extracts time-domain features through a TCN time-domain feature extractor with multi-scale receptive fields and designed contrastive loss function, combines Fast Fourier Transform (FFT) to obtain frequency-domain features, and constructs neighborhood networks for feature smoothing using quaternion distance metrics; then inputs the fused features into a support vector machine for classification. Experimental results show that the proposed method achieves a classification accuracy of 99.13% on a dataset containing 9 types of ground surfaces, approximately 9 percentage points higher than traditional methods. Furthermore, ablation experiments further verify the effectiveness of each module in the proposed method. This research provides a new technical approach for mobile robot environmental perception, with practical value for robot autonomous navigation and intelligent decision-making.

**Key Words:** Decoupled Design   Temporal Convolutional Network   Contrastive Learning   Neighborhood Smoothing   Time-Frequency Feature Fusion

## 目 录

原创性及知识产权声明 .....	I
中文摘要及关键词 .....	II
英文摘要及关键词 .....	III
一、绪论 .....	1
(一) 研究背景和意义 .....	1
(二) 国内外研究现状 .....	2
(三) 研究内容和目标 .....	5
(四) 论文的组织结构 .....	5
二、相关理论基础 .....	7
(一) 邻域网络构建以及特征平滑策略 .....	7
(二) 对比学习 .....	9
(三) 时间卷积网络 .....	10
(四) 支持向量机 .....	11
三、算法理论分析 .....	12
(一) 算法总体框架 .....	12
(二) 邻域网络构建 .....	13
(三) 时频特征提取 .....	17
(四) 特征平滑与融合 .....	21
(五) 支持向量机分类 .....	23
四、实验结果与分析 .....	25
(一) 数据集介绍 .....	25
(二) 实验设置 .....	26
(三) 对比实验 .....	29
(四) 消融实验 .....	31
(五) 结果讨论 .....	37
五、总结与展望 .....	39
(一) 本文工作总结 .....	39

## 目 录

---

(二) 未来工作展望 .....	40
参考文献 .....	42
致 谢 .....	46

## 一、绪论

### （一）研究背景和意义

随着智能设备与自主系统的广泛应用，地面介质分类作为环境感知和智能决策的基础技术，在多个领域展现出重要价值。特别是在复杂环境下，准确识别不同地面类型对移动机器人的自主导航、无人驾驶车辆的路面适应、智能假肢的动态调整等应用至关重要。复杂环境中地面材质多样、环境条件多变，这使得准确的地面介质识别成为确保系统安全运行和提高任务完成质量的关键因素。当系统能够精确区分泥土、碎石、瓷砖等不同地面，就能做出更合理的速度调整、路径规划和力度控制决策，从而显著提升整体性能和用户体验。

然而，当前地面介质分类技术在复杂环境应用中面临三大关键挑战。首先，传感器数据的高噪声性问题尤为突出。环境振动、电磁干扰以及传感器本身的不稳定性导致采集信号中包含大量噪声，这极大影响了特征提取的质量和分类的可靠性。现有算法多缺乏有效的噪声过滤机制，导致在复杂环境中分类准确率显著下降。其次，不同地面介质特征的相似性带来巨大挑战。部分地面材质（如不同类型的瓷砖或混凝土）在传感器响应上表现出高度相似的模式，而传统单一特征空间的表征能力有限，难以捕捉这些微小但关键的差异。此外，现有算法架构普遍缺乏灵活性和适应性。大多数方法采用端到端的训练模式，系统各组件高度耦合，难以针对特定应用场景进行灵活调整，也无法根据实际需求独立优化或替换某一模块，这大大限制了算法的实用性和扩展能力。

针对以上挑战，本文提出了一种基于时频域融合与邻域平滑驱动的地面介质分类算法，创新性地采用模块化设计将特征提取与分类过程解耦，为复杂环境下的地面介质分类提供了新的技术路线。具体而言，本算法通过三个关键创新解决了现有技术的痛点：首先，针对传感器数据的高噪声问题，设计了基于四元数距离度量的邻域网络构建方法和特征平滑机制，通过利用样本间的物理相似性关系，有效过滤随机噪声和异常值，显著提高了算法在复杂环境中的抗噪性和稳定性。其次，为解决地面介质特征相似性问题，算法一方面设计了多尺度 TCN 时域特征提取器并引入对比学习策略增强特征的判别能力，另一方面融合时域和频域特征实现互补增强，成功捕捉不同地面材质间的细微差异。第

三，通过模块化设计和标准化接口，使特征提取器和分类器可以独立优化和灵活替换，大大提高了系统的适应性和可扩展性，为不同应用场景的快速适配提供了可能。

本研究开创了一种新的技术思路，通过时频特征融合、对比学习和邻域平滑的创新组合，以及特征提取与分类过程的解耦设计，为复杂环境下的地面介质分类提供了高性能、高适应性的解决方案，对移动机器人环境感知、智能系统环境适应性决策等领域具有重要的理论价值和实践意义。

## （二）国内外研究现状

本节从基于外部感知的地面分类研究、基于内部感知的地面分类研究以及地面分类的特征提取与模块化系统设计三个方面对相关研究现状进行综述。

### 1. 基于外部感知技术的地面分类研究

早期的地面分类研究主要依赖外部感知技术，采用两类关键传感器：激光雷达和摄像机。

激光雷达通过发射激光脉冲并测量光的反射时间来构建三维环境模型。McDaniel 等人[1]使用地面激光雷达进行地形分类和树干识别，开发了一种基于形态学特征的算法，实现了不同地形类型的分类。Pingel 等人[2]提出了一种改进的形态学滤波方法，用于空载激光雷达数据的地形分类。虽然激光雷达在测量距离和构建大规模环境模型方面表现出色，但其捕捉地面材料微小变化的能力有限，难以实现精细的地面分类。

视觉技术是另一种广泛使用的外部感知方法。Xue 等人[3]开发了一种深度学习框架，学习地面纹理的高维流形，利用无序纹理表示和局部空间信息进行识别。Zürn 等人[4]结合声学 and 视觉模态数据进行地形分类，使用声学数据训练视觉语义分割模型，实现像素级地形分类。然而，基于视觉的方法容易受到地表覆盖物、光照变化和环境条件的影响，在某些相似外观的地形上往往难以准确分类。

此外，多模态融合方法也逐渐受到重视。DeSouza 和 Kak[5]综述了移动机器人视觉导航技术，讨论了视觉与其他传感器融合的优势。但这些多模态方法通常需要大量计算资源，难以在资源有限的机器人系统上实时运行。

综上所述，基于外部感知的地面分类方法虽然能直接获取环境信息，但容易受环境条件影响，且计算复杂度高，难以满足移动机器人实时分类的需求。本文提出的基于内部传感器的分类方法，通过直接利用机器人与地面的物理交互数据，能够更可靠地识别不同地面类型。

### 2. 基于内部感知技术的地面分类研究

近年来，研究人员提出了利用内部感知信号进行地面分类的方法，主要分为传统机器学习和深度学习两种技术路线。

传统机器学习方法依赖于手工设计的特征和经典分类器。Wu 等人[6]通过收集小型双足机器人腿部的动态地面反作用力数据，结合电机扭矩和机器人步态信息，训练了一个 SVM 地形分类器，达到了超过 90% 的分类准确率。Weiss 等人[7]使用基于振动信号的支持向量机进行地形分类，证明了内部传感器数据的有效性。然而，这些方法需要复杂的数据预处理和特征工程，使得系统开发和维护变得复杂。Lomio 等人[8]提出了一个集成机器学习方法的基线模型，由 XGBoost、全卷积网络和残差网络组成，用于评估其他深度学习模型在地面分类任务上的性能。

深度学习技术近年来在内部感知信号处理领域取得了显著进展。Vulpi 等人[9]仅依靠内部感知信号来检测地形，利用安装在自主机器人上的各种内部传感器的数据，通过深度学习技术提高了分类性能。Shah 等人[10]设计了一种基于自注意力自编码器的深度学习模型，直接从心电图（ECG）信号中学习特征表示，避免了手动特征提取的复杂性。尽管深度学习方法在特征自动提取方面具有优势，但其端到端的训练方式导致系统各组件高度耦合，难以针对不同环境或硬件平台灵活调整。

目前，基于内部感知的地面分类研究仍面临两大挑战：首先，传感器数据的高噪声性和不同地面特征的相似性使得特征提取和分类变得困难；其次，现有方法多采用端到端训练模式，系统各部件高度耦合，难以灵活调整。本文提出的基于对比学习和邻域平滑的两阶段分类算法，正是针对这些挑战设计的，通过特征学习和模块化设计提高系统的性能和灵活性。

### 3. 地面分类的特征提取与模块化系统设计研究

#### (1) 时频域特征提取方法

现有的特征提取方法主要分为时域特征提取和频域特征提取两类。Ojeda 等人[11]提出了基于惯性测量单元 (IMU) 的时域特征提取方法, 通过计算加速度和角速度信号的统计量来表征不同地面类型。Brooks 和 Iagnemma[12]利用功率谱密度 (PSD) 特征表征地面的粗糙度和刚度, 在频域分析方面取得了良好的分类效果。随着研究的深入, 时频域融合方法逐渐受到重视。Hoepflinger 等人[13]结合时域统计特征和小波变换系数, 实现了腿式机器人的地面分类。Aarabi 等人[14]提出了多通道时频特征融合框架, 使用卷积神经网络分别提取时域和频域特征, 并通过特征融合层进行整合。然而, 这些方法虽然具有更强的表达能力, 但特征间的互补性和冗余性未得到充分考虑, 且缺乏有效的特征学习机制。

### (2) 对比学习研究

对比学习作为一种新兴的表示学习方法, 近年来在计算机视觉领域取得了突破性进展。Chen[15]等人提出的 SimCLR 框架通过数据增强和对比损失学习视觉表示, 取得了显著效果。Zhou 等人[16]将对比学习应用于传感器信号分析, 证明了其在捕捉信号内在结构方面的有效性。然而, 对比学习在机器人地面介质分类中的应用研究仍处于起步阶段, 尚缺乏针对传感器时序数据特点的对比学习框架。

### (3) 模块化系统设计

在地面分类领域, Zhang 等人[17]研究了特征提取与分类器解耦的系统架构, 通过标准化特征表示实现了分类器的灵活替换。这些研究虽然在系统架构设计上有所突破, 但在特征提取器的独立优化和特征质量提升方面仍有待深入。

综上所述, 现有的地面分类特征提取和系统设计研究存在三方面不足: 一是缺乏有效的特征学习策略和噪声处理机制; 二是对比学习在地面分类中的应用研究不足; 三是模块化系统设计缺乏将特征提取和分类过程有效解耦的方法。本文提出的对比学习与邻域平滑协同的移动机器人地面介质分类算法, 通过模块化设计将特征提取与分类过程解耦, 并通过对比学习和邻域平滑提高特征质量, 旨在克服这些不足, 提高地面分类的性能和系统的灵活性。本章的相关研究工作发表在文献[18]

### （三）研究内容和目标

本研究旨在解决复杂环境下地面介质分类中的高噪声传感器数据处理和不同地面介质特征的相似性问题，提出一种时频域融合与邻域平滑驱动的地面介质分类算法。该算法采用模块化设计思想，将特征提取与分类过程解耦，通过三个核心方面的创新提高分类性能：

首先，在邻域网络构建方面，本研究设计了基于四元数距离度量的邻域识别方法，通过分析机器人方向数据的四元数表示，构建样本间的相似性关系。借助此邻域网络进行特征平滑，有效减少噪声影响，提高特征的稳定性。

其次，在特征提取方面，本研究提出了时域和频域特征的互补提取与融合方案。时域特征通过设计的多尺度 TCN 和对比学习策略提取，捕捉不同时间尺度的信息；同时，采用快速傅里叶变换（FFT）提取频域特征，捕捉传感器信号的周期性模式。这种双域特征融合极大增强了对不同地面类型的区分能力。

第三，通过将特征提取与分类过程分离，设计一种两阶段分类算法，实现系统的灵活性和高性能。研究重点关注如何设计时频特征融合策略，以及如何构建模块化接口，使特征提取器和分类器可以独立优化和灵活替换，以适应不同应用场景的需求。

### （四）论文的组织结构

在论文组织结构上，全文共分为五章，具体描述如下：

第一章介绍了面向复杂环境下的地面介质分类算法的研究背景以及本文提出算法的研究意义，并从基于外部感知技术、内部感知技术的地面分类研究以及地面分类的特征提取和模块化设计，三个角度展开文献综述。

第二章介绍了邻域网络构建以及特征平滑策略、对比学习、时间卷积网络和支持向量机的已有工作，为本文提出的算法提供理论基础。

第三章对本文提出的基于时频域融合和邻域平滑驱动的地面介质分类算法进行理论分析，阐述了每个功能模块的设计思想和原因。

第四章是实验部分，一方面，全面的介绍了实验的硬件环境以及相关的参数选择，另一方面，将本文提出的方法和领域内的其他方法进行了多个指标下的比较，最后通过消融实验验证了每个模块是否发挥作用。

第五章对本文所提出的方法进行了综合概况和提炼，阐述未来值得改进的方向。

## 二、相关理论基础

### (一) 邻域网络构建以及特征平滑策略

#### 1. 四元数表示

四元数是复数的扩展，由实部和三个虚部组成，可表示为  $q = w + xi + yj + zk$ ，其中  $w, x, y, z$  是实数， $i, j, k$  是虚部单位[19]。四元数满足以下乘法关系：

$$\begin{cases} i^2 = j^2 = k^2 = -1 \\ ijk = -1 \end{cases} \quad (2.1)$$

四元数在表示三维空间中的旋转时具有特殊优势。给定旋转轴  $\vec{v} = (v_x, v_y, v_z)$  和旋转角度  $\theta$ ，对应的四元数可表示为[20]：

$$q = \cos \frac{\theta}{2} + \sin \frac{\theta}{2} (v_x i + v_y j + v_z k) \quad (2.2)$$

相比欧拉角，四元数避免了万向节锁问题，且计算效率更高，因此在机器人方向表示中得到广泛应用。

#### 2. 距离度量

四元数之间的距离可通过以下方式计算[21]：

$$d(q_1, q_2) = \arccos(|q_1 \cdot q_2|) \quad (2.3)$$

其中  $q_1 \cdot q_2$  表示四元数的点积， $\arccos$  是反余弦函数（也写作  $\cos^{-1}$ ），它是余弦函数的反函数， $\arccos$  的作用是将单位四元数之间的点积值（范围在  $[-1, 1]$  之间）转换为对应的角度（范围在  $[0, \pi]$  之间）。

#### 3. 空间分区技术

为提高邻域搜索效率，采用网格分区技术将四元数空间划分为多维网格[22]。每个四元数坐标  $(x, y, z, w)$  映射到网格坐标：

$$pos(x) = \lfloor (x + 1) / 2 \cdot mesh \rfloor \quad (2.4)$$

其中  $mesh$  为网格尺寸。考虑到边界情况，实际实现中使用：

$$pos_2(x) = \begin{cases} \max\left(\left\lfloor \frac{(x+1)}{2} \cdot mesh - 0.5 \right\rfloor, 0\right), \\ \min\left(\left\lfloor \frac{(x+1)}{2} \cdot mesh + 0.5 \right\rfloor, mesh - 1\right) \end{cases} \quad (2.5)$$

这种设计的目的是：处理边界情况时，确保映射结果总是落在有效的网格范围内  $[0, mesh - 1]$ ，同时对于接近网格边界的点，通过 $\pm 0.5$  的偏移确保它们被正确分配到相邻的网格单元，这种网格映射策略常用于空间索引和邻域搜索算法，特别是在四元数空间中进行相似性搜索时，可以显著降低搜索复杂度。

#### 4.核密度估计

核密度估计是一种非参数概率密度估计方法，用于估计随机变量的概率分布[23]。给定样本  $\{x_1, x_2, \dots, x_n\}$ ，其核密度估计为：

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (2.6)$$

其中 $K$ 是核函数， $h$ 是带宽参数。本研究采用高斯核：

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.7)$$

带宽参数  $h$  控制估计的平滑度，较小的带宽产生更精细的分布，但可能过拟合；较大的带宽产生更平滑的分布，但可能欠拟合。

#### 5.贝叶斯概率

利用贝叶斯定理将距离度量转换为连接概率[24]：

$$P(\text{connected}|d) = \frac{P(d|\text{connected})P(\text{connected})}{P(d)} \quad (2.8)$$

对数形式为：

$$\begin{aligned} \log P(\text{connected}|d) &= \\ \log P(d|\text{connected}) - \log P(d) &+ \log P(\text{connected}) \end{aligned} \quad (2.9)$$

$$\begin{aligned} \text{prob}(d) &= \exp(\log P(d|\text{connected}) - \log P(d)) \\ &= \frac{P(d|\text{connected})}{P(d)} \end{aligned} \quad (2.10)$$

其中 $P(\text{connected})$ 是先验概率， $P(d|\text{connected})$  是似然概率， $P(d)$ 是边缘概率， $d$  表示距离度量，即两个四元数之间的距离值。在实现中，通过核密度估计计算这些概率，最终通过公式（2.10）得到连接概率。

#### 6.特征平滑

基于邻域的特征平滑通过加权平均减少噪声影响[25]。给定样本 $i$  的特征向量  $f_i$  和其邻域  $N(i)$ ，平滑后的特征为：

$$\tilde{f}_i = \frac{1}{|N(i)|} \sum_{j \in N(i)} f_j \quad (2.11)$$

其中  $\tilde{f}_i$  表示样本  $i$  的平滑后特征向量， $|N(i)|$  表示邻域集合的大小（元素数量）， $\sum_{j \in N(i)} f_j$  表示邻域内所有样本特征向量的求和。这种平滑策略在保持特征区分性的同时，有效减少噪声影响。

## (二) 对比学习

对比学习是一种无监督学习方法，目标是学习能够将相似样本拉近、不同样本推远的表示[15]。其核心思想是通过最小化对比损失函数：

$$\mathcal{L}_{contrastive} = \sum_{i=1}^N \sum_{j=1}^N y_{ij} d(f_i, f_j)^2 + (1 - y_{ij}) \max(0, m - d(f_i, f_j))^2 \quad (2.12)$$

其中  $y_{ij} = 1$  表示样本  $i$  和  $j$  相似， $y_{ij} = 0$  表示不相似， $d(f_i, f_j)$  是特征空间中的距离， $m$  是边界参数。

对比损失函数通过两项机制同时优化特征空间：对于相似样本对（ $y_{ij} = 1$ ），第一项  $y_{ij} d(f_i, f_j)^2$  惩罚它们之间的距离，迫使模型学习将相似样本映射到特征空间中的接近位置；而对于不相似样本对（ $y_{ij} = 0$ ），第二项  $(1 - y_{ij}) \max(0, m - d(f_i, f_j))^2$  仅当它们距离小于边界  $m$  时产生惩罚，确保不相似样本之间保持足够远的距离，超过边界后不再额外惩罚，从而在特征空间中形成有意义的聚类结构。

在本研究中，将对比学习应用于时域特征提取，通过设计批内对比损失[26]：

$$\mathcal{L}_{batch} = \frac{1}{B(B-1)} \sum_{i=1}^B \sum_{j=1, j \neq i}^B \mathbb{1}_{y_i=y_j} d(f_i, f_j)^2 + \mathbb{1}_{y_i \neq y_j} \max(0, m - d(f_i, f_j))^2 \quad (2.13)$$

其中  $B$  是批量大小， $\mathbb{1}$  是指示函数。这种方法显著增强了特征的判别能力，使相同地面类型的样本特征更加接近，不同类型的样本特征更加分离。

批内对比损失函数是标准对比损失的高效变体，它将计算限制在当前批次内而非整个数据集，通过指示函数  $\mathbb{1}_{y_i = y_j}$  和  $\mathbb{1}_{y_i \neq y_j}$  分别识别同类和异类样本对，同时引入归一化因子  $\frac{1}{B(B-1)}$  使损失对批量大小不敏感；这种方法不仅计算效

率高，适合随机梯度下降等迭代优化方法，而且能在训练过程中动态适应数据分布，特别适合时域特征提取，能有效捕获时间序列数据中的时态模式，提高特征判别能力和模型泛化性能。

### (三) 时间卷积网络

时间卷积网络（TCN）是一种专门处理时序数据的深度学习架构[27]。

TCN 通过扩张卷积捕捉不同时间尺度的特征，扩张卷积的定义为：

$$F(s) = \sum_{i=0}^{k-1} f(i) \cdot x(s - d \cdot i) \quad (2.14)$$

其中  $k$  是卷积核大小， $d$  是扩张率，随层数增加呈指数增长： $d_l = 2^l$ 。

TCN 的核心优势在于其能够有效捕获长期依赖关系，同时保持较低的计算复杂度。通过扩张卷积，TCN 可以指数级扩大感受野，使网络能够"看到"更长的历史序列信息[28]。一个  $L$  层的 TCN，其感受野大小可以表示为：

$$\begin{aligned} RF &= 1 + (k - 1) \sum_{i=0}^{L-1} d_i \\ &= 1 + (k - 1) \sum_{i=0}^{L-1} 2^i \\ &= 1 + (k - 1)(2^L - 1) \end{aligned} \quad (2.15)$$

TCN 不会像 RNN 那样受到梯度消失或梯度爆炸问题的影响[29]。TCN 的基本构建块通常包含残差连接和一维因果卷积。残差块可以表示为：

$$o = \text{Activation}(x + F(x)) \quad (2.16)$$

其中  $F(x)$  是残差映射函数，通常由两层扩张卷积组成。因果卷积确保了模型在预测时只使用过去和当前的信息，对于输入序列  $x = [x_1, x_2, \dots, x_T]$ ， $t$  时刻的输出  $y_t$  只依赖于  $t$  及之前的输入：

$$y_t = f(x_t, x_{t-1}, \dots, x_{t-r}) \quad (2.17)$$

TCN 的另一个显著特点是其并行计算能力。与 RNN 必须按顺序处理输入不同，TCN 可以并行处理整个序列，这显著提高了计算效率[30][31]。

研究表明，TCN 在多种时间序列任务中，如时间序列分类[32]、时间序列预测[33]和动作识别[34]等领域，都展现出优于传统 RNN 模型的性能。通过调

整扩张率和网络深度，研究人员可以灵活地控制模型的感受野大小，从而适应不同应用场景的需求。

### （四）支持向量机

支持向量机（SVM）是一种强大的监督学习算法，通过寻找最大间隔超平面分离不同类别的数据[35]。对于线性可分数据，决策函数为：

$$f(x) = \text{sign}(w^T x + b) \quad (2.18)$$

其目标是求解最优化问题：

$$\min_{w,b} \frac{1}{2} |w|^2 \quad \text{s.t. } y_i(w^T x_i + b) \geq 1, i = 1, 2, \dots, n \quad (2.19)$$

对于非线性可分数据，SVM通过核函数将数据映射到高维空间，常用的径向基核函数（RBF）定义为：

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (2.20)$$

其中 $\gamma$ 控制核函数宽度。SVM的优势在于高维空间中仍表现良好、泛化能力强，且支持调整类别权重解决样本不平衡问题。

在上述公式中， $w$ 是超平面的法向量，决定了超平面的方向； $b$ 是偏置项，决定了超平面的位置； $x$ 代表特征向量； $y_i \in \{-1, 1\}$ 是样本标签，表示样本所属的类别[36]。Sign（）函数表示符号函数，当其参数大于0时输出+1，小于0时输出-1，用于将模型的原始输出转换为类别预测[37]。

最优化问题中，目标函数 $\frac{1}{2} |w|^2$ 等价于最大化间隔 $\frac{2}{|w|}$ ，因为 $|w|$ 越小，间隔越大[38]。约束条件 $y_i(w^T x_i + b) \geq 1$ 确保所有样本都被正确分类，且距离决策边界至少有 $\frac{1}{|w|}$ 的距离。这些位于最小间隔边界上的样本点即为支持向量，它们满足 $y_i(w^T x_i + b) = 1$ ，对确定最终决策边界起着关键作用。

三、算法理论分析

(一) 算法总体框架

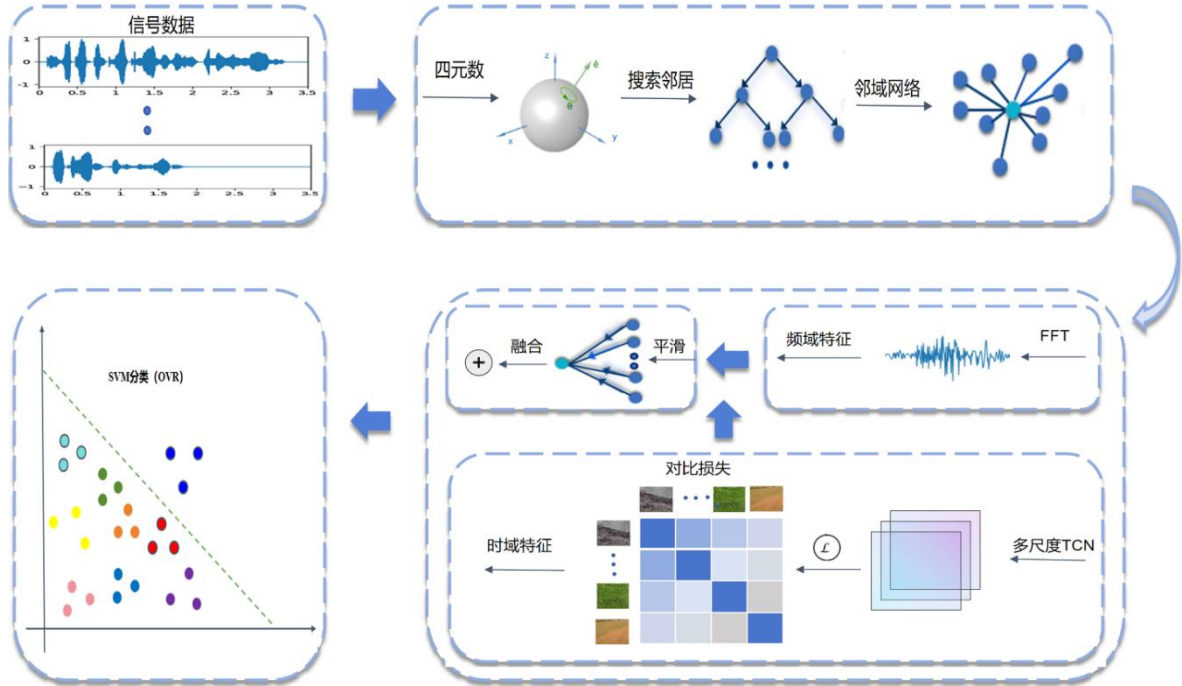


图 3.1 算法总框架

如图 3.1 所示，本算法将地面介质分类问题分解为四个关键模块：邻域网络构建、时频特征提取、特征平滑与融合、SVM 分类。算法首先基于方向四元数构建邻域网络；然后分别提取时域和频域特征，时域特征通过多尺度 TCN 和对比学习策略提取，频域特征通过 FFT 获得；接着利用邻域网络对特征进行平滑并加权融合；最后使用 SVM 进行分类。这种模块化设计使特征提取器和分类器均可独立优化和灵活替换，提高了系统适应性和可扩展性。算法的输入是机器人传感器数据，包括线性加速度、角速度和方向四元数；输出是地面介质的类别预测。

算法的核心创新点包括：（1）基于四元数距离度量的邻域网络构建并基于邻域网络进行特征平滑；（2）基于多尺度 TCN 和对比学习的时域特征提取；（4）时频特征的加权融合策略；（5）解耦特征提取与分类过程。

## (二) 邻域网络构建

邻域网络构建是算法的重要基础，其目的是识别具有相似特性的样本，为后续特征平滑提供支持。

### 1. 四元数距离度量函数

方向四元数是描述机器人姿态的有效表示方式，本算法设计了特殊的度量函数，用于评估不同四元数序列之间的相似性。

给定两个四元数序列点对 $(p_0, p_1)$ 和 $(p_2, p_3)$ ，其度量函数定义为：

$$\text{metric}(p_0, p_1, p_2, p_3) = \min(d(p_1, p_2), \max(d(2p_1 - p_0, p_2), d(2p_2 - p_3, p_1))) \quad (3.1)$$

其中， $d(p, q)$ 表示四元数 $p$ 和 $q$ 之间的距离， $2p_1 - p_0$ 表示从 $p_0$ 到 $p_1$ 的延伸预测轨迹。公式(3.1)考虑了直接距离和考虑运动方向的距离，能够更准确地描述四元数序列的相似性。

对于两个序列的端点评估，定义距离函数：

$$\text{dist}(a_0, a, b, b_0) = \max(-a \cdot b, \text{metric}(a_0, a, b, b_0)) \quad (3.2)$$

其中， $a_0$ 和 $a$ 是第一个序列的末尾两个点， $b$ 和 $b_0$ 是第二个序列的开始两个点， $a \cdot b$ 表示四元数的点积。公式(3.2)结合了方向相似性和序列连续性，为邻域识别提供了全面的度量标准。

### 2. 基于四元数距离度量的初始邻域网络构建

基于上述距离度量函数，算法 3.1 描述了初始邻域网络构建过程。为提高效率，我们采用空间分区技术，将四元数映射到网格中，实现快速邻域搜索，将算法复杂度从 $O(N^2)$ 降至近似 $O(N \cdot \log N)$ 。

在算法开始前，我们先定义以下空间分区函数（ $mesh = 100$ ）：

$$\text{pos}(x) = \lfloor (x + 1) \cdot \text{mesh} / 2 \rfloor$$

$$\text{pos2}(x) = \{ \max(\lfloor (x + 1) \cdot \text{mesh} / 2 - 0.5 \rfloor, 0), \min(\lfloor (x + 1) \cdot \text{mesh} / 2 + 0.5 \rfloor, \text{mesh} - 1) \}$$

$$\text{ibucket}(q) = (\text{pos}(q.x), \text{pos}(q.y), \text{pos}(q.z), \text{pos}(q.w))$$

$$\text{gbuckets}(q) = \{(i, j, k, l) | i \in \text{pos2}(q.x), j \in \text{pos2}(q.y), k \in \text{pos2}(q.z), l \in \text{pos2}(q.w)\}$$

各个函数的说明如下：

**pos (x) 函数：**将四元数的每个分量从[-1,1]的范围映射到[0,mesh-1]的整数索引。这是一个基本的坐标量化函数，通过线性映射和取整，将连续空间转换为离散网格单元。

**pos2 (x) 函数：**是 pos (x) 的扩展版本，不只返回单个网格索引，而是返回原点周围的两个相邻索引。这允许我们考虑边界情况，确保靠近网格边缘的四元数也能被正确索引和检索。这种"模糊匹配"机制提高了空间分区的容错性。

**ibucket (q) 函数：**将整个四元数 q 映射到四维网格中的唯一单元坐标。它通过对四元数的每个分量 (x,y,z,w) 应用 pos 函数，生成一个四维网格坐标，从而实现四元数到网格单元的精确映射。

**gbuckets (q) 函数：**生成四元数 q 周围的所有潜在网格单元集合。它使用 pos2 函数获取每个分量可能的网格索引，然后计算它们的笛卡尔积，形成一组邻近的网格单元坐标。这使得算法能够检查不仅包含目标四元数的网格单元，还包括其周围的网格单元，从而提高了搜索的完整性。

通过算法 3.1 构建了一个有向图结构，其中节点表示样本，边表示样本间的相似关系。这种邻域网络实际上是对原始数据内在结构的提取，将样本间的高维复杂关系简化为二元连接关系，在保留关键信息的同时，降低了后续处理的复杂度。特别是，算法通过同时保存前向和后向连接，构建了一个完整的邻域拓扑结构，为后续的特征平滑奠定了基础。

### 3.基于核密度估计的连接概率计算

为量化序列之间的连接可能性，本算法使用核密度估计 (KDE) 计算连接概率。首先，计算两个概率分布：

已知是连接的序列对之间的距离分布 (使用同一序列内部的点)

所有可能的序列对之间的距离分布

对于第一个分布，使用前面收集的 $\mathcal{D}$ ：

$$p_{\text{connected}}(x) = \text{KDE}_{\mathcal{D}}(x) + \delta_0 \quad (3.3)$$

其中， $\text{KDE}_{\mathcal{D}}$ 表示在数据集 $\mathcal{D}$ 上训练的核密度估计模型， $\delta_0 = -\log(N)$ 为先验概率调整项， $N$ 为样本数量。

对于第二个分布，使用初始邻域网络中的所有距离：

$$\mathcal{D}_2\{d|(j, d) \in \text{neigh\_fwd}[i], i \in \{1, 2, \dots, N\}\} \quad (3.4)$$

### 三、算法理论分析

$$p_{\text{all}}(x) = \text{KDE}_{\mathcal{D}_2}(x) + \delta \quad (3.5)$$

其中， $\delta = \log(|\mathcal{D}_2|/(N^2))$ 为调整因子，考虑了观察到的连接相对于可能连接总数的比例。

---

算法 3.1: 基于四元数距离度量的初始邻域网络构建

---

输入: 样本四元数序列集合  $Q = \{q_1, q_2, \dots, q_n\}$ , 其中  $q_i$  是长度为 128 的四元数序列

输出: 初始邻域网络  $neigh\_fwd, neigh\_rev$

```
1 // 计算距离分布和阈值
2  $D \leftarrow \emptyset$ 
3   for each  $e \in Q$  do
4     for  $i \in \{0, 16, 32, \dots, 112\}$  do
5        $d \leftarrow \text{metric}(e_i, e_{i+1}, e_{i+2}, e_{i+3})$ 
6        $D \leftarrow D \cup \{d\}$ 
7     end for
8   end for
9  $\text{thresh} \leftarrow \text{quantile}(D, 0.999)$ 
10 // 构建空间索引
11  $\text{ind}[0] \leftarrow \{\}$  // 存储序列起点
12  $\text{ind}[1] \leftarrow \{\}$  // 存储序列终点
13 for  $i \leftarrow 1$  to  $N$  do
14   for  $j \in \{0, 1\}$  do
15      $k \leftarrow \text{ibucket}(q_i[j \cdot 127])$ 
16     if  $k \notin \text{ind}[j]$  then  $\text{ind}[j][k] \leftarrow \{i\}$ 
17     else  $\text{ind}[j][k] \leftarrow \text{ind}[j][k] \cup \{i\}$ 
18   end for
19 end for
20 // 构建初始邻域网络
21  $neigh\_fwd \leftarrow [\emptyset, \emptyset, \dots, \emptyset]$  // 大小为  $N$  的空列表数组
22  $neigh\_rev \leftarrow [\emptyset, \emptyset, \dots, \emptyset]$  // 大小为  $N$  的空列表数组
23 for  $i \leftarrow 1$  to  $N$  do
24   for each  $k \in \text{gbuckets}(q_i[127])$  do
25     if  $k \in \text{ind}[0]$  then
26       for each  $j \in \text{ind}[0][k]$  where  $j \neq i$  do
27          $d \leftarrow \text{dist}(q_i[126], q_i[127], q_j[0], q_j[1])$ 
28         if  $d < \text{thresh}$  then \
29            $neigh\_fwd[i] \leftarrow neigh\_fwd[i] \cup \{(j, d)\}$  //  $i \rightarrow j$ 
30            $neigh\_rev[j] \leftarrow neigh\_rev[j] \cup \{(i, d)\}$  //  $j \leftarrow i$ 
31         end if
32       end for
```

---

```

33     end if
34   end for
35 end for
36 return neigh_fwd,neigh_rev

```

最后，使用贝叶斯公式计算给定距离的连接概率：

$$\begin{aligned}
 P(\text{connected}|x) &= \frac{P(x|\text{connected}) \cdot P(\text{connected})}{P(x)} \\
 &= \exp(p_{\text{connected}}(x) - p_{\text{all}}(x))
 \end{aligned} \tag{3.6}$$

公式（3.6）表示在观察到距离为 $x$ 的情况下，两个序列真正连接在一起的后验概率。然后将初始邻域网络中的距离转换为连接概率，并按概率降序排序。

#### 4.基于连接概率的最终邻域网络构建

基于计算的连接概率，构建最终的邻域网络，控制邻域大小以确保网络的稀疏性。算法 3.2 采用贪心搜索策略构建最终邻域网络的设计理念深植于图论与优化理论。该算法通过基于连接概率的双向搜索机制，在前向和后向两个方向同时探索可能的邻居节点，确保发现最优连接路径；实现了概率驱动的邻域构建，通过累积概率评估整条路径的可靠性，而非仅考虑单跳距离；引入了概率阈值  $p\_cutoff$  和邻域大小上限  $len\_cutoff$  的双重控制机制，在保证邻域质量的同时维持网络稀疏性，有效避免了低质量连接带来的噪声影响；算法的设计体现了"局部最优"与"全局约束"的平衡，通过在每个迭代中选择最高概率的邻居并考虑累积效应，形成一种介于最短路径与最大流问题之间的特殊图论优化问题。这种基于概率的贪心邻域构建策略不仅降低了算法复杂度，更为后续特征平滑提供了高质量的邻域结构，显著提升了地面介质分类的鲁棒性和准确性。

算法 3.2: 基于连接概率的最终邻域网络构建

输入：前向和后向邻居列表  $neigh\_fwd, neigh\_rev$ （已转换为概率形式）

参数：概率阈值  $p\_cutoff = 0.5$ ，邻域大小上限  $len\_cutoff = 30$

输出：最终邻域网络  $N = N_1, N_2, \dots, N_n$

```

1  neigh ← [neigh_fwd,neigh_rev]
2  N ← [∅,∅, ..., ∅] // 大小为 N 的空列表数组
3  for i ← 1 to i do
4    visited ← {i};visited_o ← [i] //已访问集合和有序列表
5    front ← [i,i];front_p ← [1.0,1.0] //搜索前沿和概率
6    while |visited| < len_cutoff do
7      p_max ← -∞; n_max ← None; d_max ← None
8      for d ∈ {0,1} do //检查前向和后向方向

```

### 三、算法理论分析

---

```
9       for each (n_idx, n_prob) ∈ neigh[d][front[d]] do
10          if n_idx ∉ visited and n_prob · front_p[d] > p_max then
11             p_max ← n_prob · front_p[d]; n_max ← n_idx; d_max ← d
12          end if
13       end for
```

---

续算法 3.2: 基于连接概率的最终邻域网络构建

---

```
14       end for
15       if n_max = None or n_max < p_cutoff then break end if
16       visited ← visited ∪ {n_max}; visited_o.append(n_max)
17       front[d_max] ← n_max; front_p[d_max] ← p_max
18     end while
19     N[i] ← visited_o
20   end for
21   return N
```

---

#### 5. 邻域网络特性

**相似性保证:** 邻域中的样本在四元数空间中具有高度相似性, 保证了平滑操作的有效性。

**稀疏连接:** 每个样本最多有 `len_cutoff` 个邻居, 保证了网络的稀疏性和计算效率。

**概率阈值:** 通过概率阈值 `p_cutoff` 控制邻居的质量, 确保只保留高可信度的连接。

**空间局部性:** 空间分区技术确保只考虑可能的邻居, 大大提高了算法效率。

#### (三) 时频特征提取

##### 1. 时域特征提取

###### (1) 多尺度 TCN 架构设计

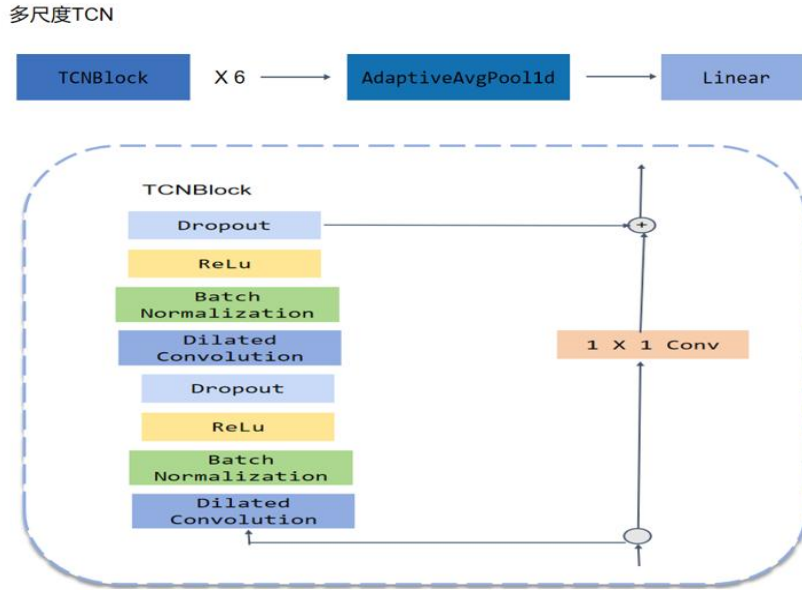


图 3.2 多尺度 TCN

图 3.2 展示了本文设计的多尺度 TCN 特征提取器架构，多尺度 TCN 设计的理论基础源于感受野理论[39]，这一概念最初来自视觉神经科学，在时序分析中有着深远意义。本研究设计的多尺度 TCN 通过扩张系数呈指数增长（ $2^0$ 至 $2^5$ ）的扩张卷积层次结构，形成了理论感受野大小为 127 的网络架构，几乎覆盖整个输入序列。这种设计精确匹配了地面介质分类的本质需求：地面介质的识别依赖于机器人传感器捕获的多尺度振动模式，既需要捕捉局部短时特征（如瞬时冲击），也需要理解长程依赖关系（如持续振动频率）。当机器人从一种地面过渡到另一种地面时，传感器信号会呈现出独特的时序模式，这些模式在不同时间尺度上表现出不同特征。多尺度 TCN 的每一层通过不同扩张率捕捉不同时间跨度的依赖关系，浅层网络（小扩张率）提取局部快速变化特征，深层网络（大扩张率）则识别长期持续的振动模式，这种层次化感受野结构使网络能够同时关注微观细节和宏观模式，为地面介质的精确分类提供了全面的时序特征表示，从理论上解决了传统单一尺度方法难以平衡短期与长期依赖性的根本问题。

### (2) TCN 特征提取器的构成

**TCN 基本块：**每个 TCN 块包含两个扩张卷积层，残差连接和批量归一化，定义如下：

$$\left\{ \begin{array}{l} z^{(1)} = \text{Conv1D}(x; \theta_{conv1}, k, d) \\ z^{(2)} = \text{BatchNorm}(z^{(1)}) \\ z^{(3)} = \text{ReLU}(z^{(2)}) \\ z^{(4)} = \text{Conv1D}(z^{(3)}; \theta_{conv2}, k, d) \\ z^{(5)} = \text{BatchNorm}(z^{(4)}) \\ \text{TCNBlock}(x) = \text{ReLU}(z^{(5)} + \text{Downsample}(x)) \end{array} \right. \quad (3.7)$$

其中， $k$ 是卷积核大小， $d$ 是扩张率， $\text{Downsample}(x)$ 是当输入输出通道数不同时用于调整维度的  $1 \times 1$  卷积。

**多尺度层次结构：**由 6 个 TCN 基本块组成，每层的扩张率呈指数增长，分别为 1、2、4、8、16 和 32（即  $2^0$  到  $2^5$ ），使模型能够捕捉不同时间尺度上的特征。每个卷积层使用大小为 3 的卷积核，隐藏单元数设置为 64。该模型接收 6 个输入通道，对应 6 种不同的传感器信号，最终输出维度为 32 的特征表示，有效地提取和融合了多种传感器数据在不同时间尺度上的信息。

$$RF = 1 + (k - 1) \times \sum_{i=0}^{L-1} d_i \quad (3.8)$$

TCN 的有效感受野大小可以通过公式（3.8）计算，其中 $k$ 是卷积核大小， $d_i$ 是第 $i$ 层的扩张率， $L$ 是层数。在我们的设计中，扩张率呈指数增长，使 6 层 TCN 的有效感受野达到  $1 + 2 \times (1 + 2 + 4 + 8 + 16 + 32) = 1 + 2 \times 63 = 127$ ，几乎覆盖了整个输入序列长度（128），这使得网络能够捕获从短期到长期的时间依赖关系。卷积核大小选择为 3 是计算复杂度和表达能力的折中，隐藏单元数 64 则提供了足够的模型容量而不致过拟合。

**输出层：**输出层由全局平均池化层和全连接层组成，其中全局平均池化将不同长度的时序特征压缩为固定长度的表示，有效消除了序列长度的影响；随后，全连接层将池化后的特征进一步映射到最终的特征空间，实现了从时序特征到固定维度特征表示的转换，为后续任务提供了统一的特征输入。

### （3）对比损失函数

为提高特征的区分能力，我们设计了基于欧氏距离的对比损失函数，定义如下：

$$\mathcal{L}_{contrast}(F, Y, margin) = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \left[ \mathbf{1}_{y_i=y_j} \cdot d_{ij}^2 + \mathbf{1}_{y_i \neq y_j} \cdot \max(0, margin - d_{ij})^2 \right] \quad (3.9)$$

其中对于每个批次的的数据，对于每个批次的的数据， $F = \{f^1, f^2, \dots, f_n\}$ 表示批次中的特征向量集合， $Y = \{y^1, y^2, \dots, y_n\}$ 为对应的类别标签。特征向量之间的相似度通过欧氏距离  $d_{ij} = \|f_i - f_j\|^2$  来度量。模型设定了  $margin = 1.0$  作为不同类别样本之间应保持的最小距离阈值，以确保类间分离。此外， $\mathbf{1}_{condition}$  作为指示函数，当条件满足时取值为 1，否则为 0，用于在损失函数中选择性地对样本对进行惩罚或奖励，从而优化特征空间中的样本分布。表示批次中的特征向量集合， $Y = \{y^1, y^2, \dots, y_n\}$ 为对应的类别标签。特征向量之间的相似度通过欧氏距离  $d_{ij} = \|f_i - f_j\|^2$  来度量。模型设定了  $margin=1.0$  作为不同类别样本之间应保持的最小距离阈值，以确保类间分离。此外， $\mathbf{1}_{condition}$  作为指示函数，当条件满足时取值为 1，否则为 0，用于在损失函数中选择性地对样本对进行惩罚或奖励，从而优化特征空间中的样本分布。

该损失函数本质上是同类样本在特征空间中聚集，同时将不同类别样本分开。具体而言，对于同类样本对  $(i, j)$ ，损失项  $d_{ij}^2$  直接惩罚它们之间的欧氏距离，促使特征靠近；对于不同类别样本对，损失项  $\max(0, margin - d_{ij})^2$  只有在距离小于  $margin$  时才产生惩罚，这种铰链损失（hinge loss）形式鼓励不同类别样本间至少保持  $margin$  的距离，而不会过度惩罚已经充分分开的样本，从而更高效地优化特征空间分布。

相比于传统的分类损失函数（如交叉熵），对比损失直接优化特征之间的关系而非通过分类层间接优化，这使得学习到的特征具有更好的区分性和泛化能力，特别是在后续用于特征平滑和 SVM 分类时表现更佳。

## 2. 频域特征提取

频域特征提取利用快速傅里叶变换（FFT）捕捉传感器信号的周期性模式。对于每个传感器信号（线性加速度和角速度的 x、y、z 分量），FFT 过程如下：

首先，计算 FFT 并提取前 64 个复数分量（除直流分量外）：

$$X_f[i] = \frac{1}{\sqrt{L}} \sum_{n=0}^{L-1} x_f[n] e^{-j2\pi n i / L}, \quad i = 1, 2, \dots, 64 \quad (3.10)$$

其中， $x_f[n]$  为特征  $f$  的时序信号， $L = 128$  为序列长度， $X_f[i]$  为 FFT 结果

的第 $i$ 个分量， $f \in \{ax, ay, az, lx, ly, lz\}$ 表示角速度和线性加速度的6个分量。除以 $\sqrt{L}$ 进行归一化处理。

然后，计算每个频率分量的平均幅度，用于归一化：

$$\text{norm}_f[j] = \frac{1}{N} \sum_{i=1}^N |X_f^i[j]|, \quad j = 1, 2, \dots, 64 \quad (3.11)$$

其中， $N$ 为样本数量， $X_f^i[j]$ 为第 $i$ 个样本特征 $f$ 的FFT第 $j$ 个分量， $\text{norm}_f[j]$ 为频率 $j$ 的平均幅度。

接下来，对FFT结果进行归一化处理：

$$\widehat{X}_f^i[j] = \frac{X_f^i[j]}{\text{norm}_f[j]}, \quad i = 1, 2, \dots, N, j = 1, 2, \dots, 64 \quad (3.12)$$

其中， $\widehat{X}_f^i[j]$ 为归一化后的FFT结果。

最后，提取频域特征向量：

$$V_f^i = \{ \left[ \left[ \widehat{X}_f^i[2j] + \widehat{X}_f^i[2j + 1] \right] \cdot \text{scale}_f^i \mid j = 0, 1, \dots, 31 \right\} \quad (3.13)$$

$$\text{scale}_f^i = \frac{1}{2 \cdot \text{avg}_f^i}, \quad \text{avg}_f^i = \frac{1}{64} \sum_{j=1}^{64} |\widehat{X}_f^i[j]| \quad (3.14)$$

其中， $V_f^i$ 为第 $i$ 个样本特征 $f$ 的频域特征向量（维度为32）， $\text{scale}_f^i$ 为缩放因子， $\text{avg}_f^i$ 为平均幅度。公式（3.14）通过将相邻的频率分量配对并取平均，降低了特征维度，同时保留了频谱的关键信息。

最终的频域特征通过连接所有传感器信号的特征向量获得：

$$F_{\text{freq}}^i = [V_{ax}^i, V_{ay}^i, V_{az}^i, V_{lx}^i, V_{ly}^i, V_{lz}^i] \quad (3.15)$$

其中， $F_{\text{freq}}^i \in R^{192}$ 为第 $i$ 个样本的频域特征向量，维度为 $6 \times 32 = 192$ 。

#### （四）特征平滑与融合

特征平滑与融合是算法的关键环节，通过邻域平滑减少噪声影响，通过加权融合充分利用时频域互补信息。

##### 1. 基于邻域网络的特征平滑

特征平滑利用已构建的邻域网络，对时域和频域特征分别进行平均处理，减少噪声影响，提高特征的稳定性。对于任意特征向量 $F$ ，平滑操作定义为：

$$F'_i = \frac{1}{|\mathcal{N}[i]|} \sum_{j \in \mathcal{N}[i]} F_j \quad (3.16)$$

其中， $F_i$ 为第*i*个样本的原始特征向量， $F'_i$ 为平滑后的特征向量， $\mathcal{N}[i]$ 为样本*i*的邻域， $|\mathcal{N}[i]|$ 为邻域大小。

分别对时域和频域特征应用平滑操作：

$$\begin{cases} F'_{\text{time}} &= \text{Smooth}(F_{\text{time}}, \mathcal{N}) \\ F'_{\text{freq}} &= \text{Smooth}(F_{\text{freq}}, \mathcal{N}) \end{cases} \quad (3.17)$$

公式 (3.16) - (3.17) 通过对邻域内所有样本特征取平均，有效减少了单个样本的噪声影响，提高了特征的稳定性和鲁棒性。

这种平滑机制的理论基础源于图信号处理中的拉普拉斯平滑理论[40]，它假设相似样本在特征空间中应该具有相似的表示，通过局部平均可以过滤掉高频噪声同时保留信号的基本结构。在我们的场景中，由于相邻时间点的传感器数据在物理上具有连续性，邻域平滑不仅能减少随机噪声，还能增强特征的时间一致性，尤其对于地面过渡区域和传感器临时干扰的情况尤为有效。

邻域平滑既应用于时域特征，也应用于频域特征，能够有效减少噪声并增强特征的稳定性。实验表明，这种平滑策略比单纯的特征增强或数据增强更有效，因为它充分利用了样本间的物理关联性，而不仅仅是统计相似性。

## 2.时频特征归一化与融合

为了平衡不同特征的贡献，对时域和频域特征进行归一化处理，再进行加权融合：

$$\begin{cases} F'_{\text{time}} &= \text{Smooth}(F_{\text{time}}, \mathcal{N}) \\ F'_{\text{freq}} &= \text{Smooth}(F_{\text{freq}}, \mathcal{N}) \end{cases} \quad (3.18)$$

$$F_{\text{hybrid}} = [w_{\text{time}} \cdot F_{\text{time}}^{\text{norm}}, w_{\text{freq}} \cdot F_{\text{freq}}^{\text{norm}}] \quad (3.19)$$

其中， $F'_{\text{time}}$ 和 $F'_{\text{freq}}$ 分别为平滑后的时域和频域特征， $F_{\text{time}}^{\text{norm}}$ 和 $F_{\text{freq}}^{\text{norm}}$ 为归一化后的特征， $w_{\text{time}}$ 和 $w_{\text{freq}}$ 为权重参数， $\epsilon$ 为小常数( $10^{-8}$ )，防止除零错误， $[\cdot, \cdot]$ 表示特征连接操作。

公式 (3.18) 对特征进行 $L_2$ 归一化处理，确保不同特征的数值范围相似；公式 (3.19) 通过加权连接融合时频特征，充分利用互补信息。

## (五) 支持向量机分类

### 1. 类别权重计算

为针对不同地面类型的分布不平衡问题，我们根据各类别的目标频率分布计算权重：

$$weight_i = \frac{tf_i \times n}{n_i} \quad (3.20)$$

其中： $tf_i$  为类别  $i$  的目标频率， $n$  为样本总数， $n_i$  为类别  $i$  的样本数

针对目标场景中不同地面类型的分布特性，我们设计了相应的类别权重计算机制，使 SVM 分类器能够更好地处理类别不平衡问题。

### 2. SVM 分类器设计

本算法采用支持向量机作为分类器，其决策函数定义为：

$$f(x) = \text{sign} \left( \sum_{i=1}^{N_s} \alpha_i y_i K(x_i, x) + b \right) \quad (3.21)$$

其中， $N_s$  为支持向量的数量， $\alpha_i$  为拉格朗日乘子， $y_i$  为支持向量  $x_i$  的类别标签， $K(x_i, x)$  为核函数， $b$  为偏置项。

对于多类别分类，采用“一对多”（OVR）策略：

$$f_c(x) = \sum_{i=1}^{N_s^c} \alpha_i^c y_i^c K(x_i^c, x) + b^c, \quad \forall c \in \{1, 2, \dots, C\} \quad (3.22)$$

$$\text{class}(x) = \arg \max_c f_c(x) \quad (3.23)$$

其中， $C$  为类别数量， $f_c(x)$  为第  $c$  类的决策函数， $N_s^c$  表示第  $c$  类的支持向量数量， $\alpha_i^c$  表示对应的拉格朗日乘子， $y_i^c$  表示类别标签， $x_i^c$  表示支持向量。

本算法选择径向基函数（RBF）作为核函数：

$$K(x_i, x) = \exp(-\gamma |x_i - x|^2) \quad (3.24)$$

其中， $\gamma$  为核函数参数，控制决策边界的平滑程度

支持向量机（SVM）作为本研究分类器的选择，与融合后的特征空间具有深度的理论契合性。首先，SVM 通过寻找最大间隔超平面的核心机制特别适合处理经过时频融合与邻域平滑后形成的高度结构化特征空间。这种特征空间经过对比学习和平滑处理，已使同类样本聚集、异类样本分离，为 SVM 的最大间隔原理提供了理想的数据分布基础。其次，SVM 的核函数机制（特别是 RBF

### 三、算法理论分析

---

核)能够有效处理融合特征的非线性关系,将时域与频域信息在高维空间中进行隐式映射,捕捉地面介质复杂的振动模式与时频特性。第三,SVM对高维特征的处理能力使其能够充分利用融合特征中的判别信息而不受维数灾难影响,这对融合了192维频域特征和32维时域特征的混合空间尤为重要。此外,本研究引入的基于目标频率的类别权重计算方法与SVM的支持向量理论相辅相成,共同解决了样本不平衡问题,使边缘类别(如硬瓷砖)获得合理的决策空间。最后,SVM作为一种解耦的分类器,与本研究模块化设计理念高度一致,允许特征提取器和分类器独立优化,大大提高了系统的灵活性与适应性。

## 四、实验结果与分析

### (一) 数据集介绍

本研究采用 Lomio 等人[8]于 2019 年提出的室内地面类型分类数据集。该数据集包含超过 7600 个带标注的时间序列样本，由移动机器人搭载的惯性测量单元 (IMU) 在 9 种不同地面材质上记录的传感器数据构成。数据集中的每个样本在 10 个特征上记录了 128 个连续时间点的测量值，全面捕捉了机器人在不同地面材质上运动时的动态特性。该数据集是室内地面分类领域最全面的资源之一，被广泛应用于室内导航和地面识别研究。为确保与其他比较方法的一致性，本研究选取了数据集中的前 3810 个样本进行实验。

表 4.1 详细描述了数据集中的传感器特征及其含义。这些特征可分为三类：方向四元数、角速度和线性加速度，共 10 个维度。其中，方向四元数

(Orientation) 通过四个分量 (X, Y, Z, W) 表示机器人相对于起始位置的姿态；角速度 (Angular Velocity) 通过三个分量 (X, Y, Z) 描述机器人围绕其局部坐标轴的旋转速度；线性加速度 (Linear Acceleration) 同样通过三个分量 (X, Y, Z) 表示机器人在局部坐标系中的加速度，单位为  $m/sec^2$ 。此外，Series ID 用于将输入数据与输出地面类型关联。

表 4.1 机器人传感器特征说明

特征名称	描述
方向四元数 (X, Y, Z, W)	机器人相对于起始位置的姿态，以四元数表示
角速度 (X, Y, Z)	机器人围绕其局部坐标轴的旋转速度
线性加速度 (X, Y, Z)	机器人在局部坐标系中的加速度，单位为 $m/sec^2$
Series ID	用于匹配训练集中输入和输出的组号

数据集中的目标变量 "surface" 表示机器人下方的地面类型，是模型预测的目标。数据集包含 9 种不同的地面类型，每种类型的样本比例如图 4.1 所示。从图中可以看出，样本分布存在一定不平衡，其中 "混凝土" (concrete) 比例最高，占 20.4%；"硬瓷砖" (hard\_tiles) 比例最低，仅占 0.6%。

## 四、实验结果与分析

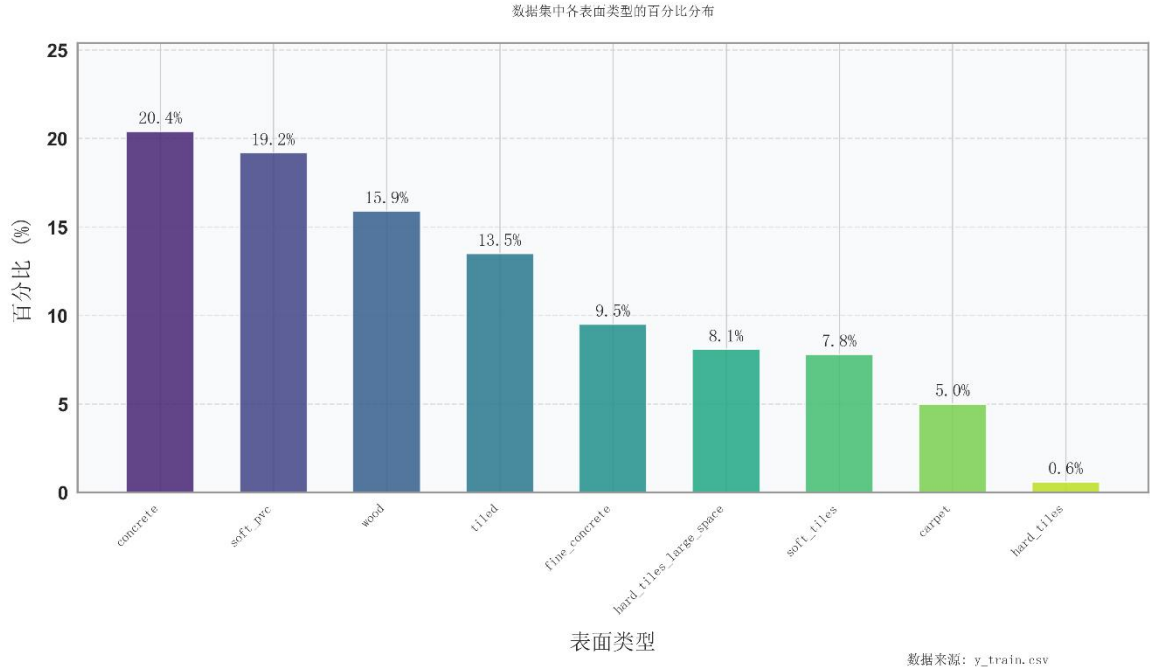


图 4.1 数据集中各地面类型分布情况

### (二) 实验设置

#### 1. 实验实现

本实验基于 PyTorch 框架实现，使用 Adam 优化器进行训练，共运行 250 个训练周期 (epoch)，初始学习率设为 0.001，并采用余弦退火策略[60]进行动态调整。具体而言，学习率按余弦函数从初始值逐渐衰减至零，随后重置为原始值，每个调参周期包含 128 个训练周期。实验设置批量大小 (batch size) 为 32，所有训练均在 NVIDIA GeForce RTX 3090 笔记本电脑 GPU 上完成。

#### 2. 评估指标

##### (1) 准确率 (Accuracy)

准确率衡量模型正确分类的样本占总样本的比例，可表示为：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

其中，TP (True Positive) 表示真正例，TN (True Negative) 表示真负例，FP (False Positive) 表示假正例，FN (False Negative) 表示假负例。

##### (2) AUC Score

## 四、实验结果与分析

AUC (Area Under the ROC Curve) 表示受试者工作特征曲线下的面积, 评估模型区分不同类别的能力。数学上可表示为:

$$AUC = \int_0^1 TPR(FPR^{-1}(t))dt \quad (4.2)$$

其中, TPR (True Positive Rate) 为真正例率, 即召回率; FPR (False Positive Rate) 为假正例率。AUC 数值范围为[0,1], 值越大表示模型性能越好。

### (3) 精确率 (Precision)

精确率衡量在所有被预测为正例的样本中, 真正例的比例:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4.3)$$

精确率反映了模型预测正例的准确性, 值越高表示假正例越少。

### (4) 召回率 (Recall)

召回率, 也称为敏感度 (Sensitivity), 衡量在所有实际为正例的样本中, 被正确识别的比例:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4.4)$$

召回率反映了模型捕获所有正例的能力, 值越高表示漏判的正例越少。

### (5) F1 分数 (F1 Score)

F1 分数是精确率和召回率的调和平均值, 平衡了两者之间的权衡:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.5)$$

F1 分数特别适用于类别不平衡的数据集, 综合考虑了精确率和召回率的表现, 取值范围为[0,1], 越接近 1 表示模型性能越好。

通过这五项指标的综合分析, 可全面评估模型在不同方面的性能表现, 为后续算法改进提供依据。

## 3. 实验流程

每组实验均采用控制变量法, 使用不同随机种子重复 5 次, 计算性能指标的平均值和方差以确保结果的可靠性。对于每次实验, 我们保持相同的数据划分比例和预处理步骤, 仅改变待研究的变量。在训练过程中, 采用早停策略 (early stopping) 以防止过拟合, 当验证集性能连续 5 个轮次未提升时停止训练。

## 四、实验结果与分析

### 4.参数设置

**特征平滑与融合模块参数：**关键参数设置详见表 4.2。

表 4.2 特征平滑与融合模块参数

参数名	描述	取值
mesh	空间分区网格大小	100
thresh	距离阈值	所有距离的 99.9%分位数
bandwidth	核密度估计带宽	0.0005
p_cutoff	概率阈值	0.5
len_cutoff	邻居数量上限	30
time_weight	时域特征权重	0.3
fre_weight	频域特征权重	0.7

**TCN 特征提取器参数：**多尺度 TCN 特征提取器采用了 6 层 TCN 结构，每层包含双卷积模块和残差连接。核大小设置为 3，隐藏通道数为 64，dropout 率为 0.2 以防止过拟合。扩张率从第一层的 1 ( $2^0$ ) 递增到第六层的 32 ( $2^5$ )，使得感受野随层数指数增长，能够捕捉不同时间尺度上的特征模式。全局池化后，通过一个全连接层将特征降维至 32 维（具体可见表 4.3）。

表 4.3 TCN 和对比学校参数

参数名	描述	取值
input_channels	输入通道数	6
hidden_units	TCN 块的隐藏单元数	64
kernel_size	卷积核大小	3
num_levels	TCN 层数	6
output_features	输出特征维度	32
margin	对比损失边界参数	1.0
input_channels	输入通道数	6
hidden_units	TCN 块的隐藏单元数	64

**对比学习参数：**对比学习采用余弦相似度作为度量函数，边缘参数设置为 1.0，批大小为 32。训练过程中使用了在线困难样本挖掘策略，每批次中选择最困难的 16 对样本进行梯度更新，有效提高了模型的判别能力（具体可见表 4.3）。

## 四、实验结果与分析

**SVM 分类器参数：**SVM 分类器采用 RBF 核函数，正则化参数 C 设置为 1.2，gamma 参数采用'scale'自适应设置。为处理类别不平衡问题，我们实现了基于目标频率的类别权重计算方法，使分类器对样本较少的类别（如硬瓷砖）具有更高的敏感度。决策函数采用 one-vs-rest 策略，更适合多类别分类任务。缓存大小设置为 2000MB 以加速训练过程（具体可见表 4.4）。

表 4.4 SVM 分类器取值

参数名	描述	取值
C	正则化参数，控制误分类惩罚力度	1.2
kernel	核函数类型	'rbf'
gamma	核函数参数	'scale'
probability	是否输出概率估计	True
decision_function_shape	决策函数形状	'ovr'（一对多）
cache_size	核矩阵缓存大小（MB）	2000

### （三）对比实验

为验证所提方法的有效性，我们将其与 16 种代表性的地面分类方法进行了对比。表 4.4 展示了各方法在 5 次重复实验中的平均性能及标准差。

表 4.4 不同方法在地面类型分类任务上的性能对比

对比方法	准确率 (%)	AUC Score (%)	精确率 (%)	召回率 (%)	F1 Score (%)
LSTM+1d-					
AX+Channel-weighting <sup>1</sup>	33.88 ±0.125	74.94 ±0.03	28.01 ±0.04	26.27 ±0.03	25.07 ±0.04
LSTM+Attention <sup>1</sup>	42.31 ±0.03	77.72 ±0.008	38.99 ±0.06	37.24 ±0.04	37.49 ±0.04
TCN	48.37 ±0.93	84.85 ±0.81	45.45 ±0.92	41.12 ±0.92	40.56 ±0.90
SVM+Feature-engineering <sup>1</sup>	49.46 ±1.49	87.58 ±0.41	44.58 ±2.45	42.28 ±1.23	31.78 ±2.46
DRCNN	55.18 ±0.54	86.48 ±0.08	52.41 ±0.16	50.52 ±0.14	50.81 ±0.13
Transformer	58.27 ±0.90	89.36 ±0.47	56.45 ±1.85	54.23 ±0.85	53.56 ±0.79
XGBoost	59.54 ±5.57	89.59 ±1.51	50.78 ±3.45	50.45 ±2.65	50.60 ±2.54
FCN	62.69 ±6.74	90.45 ±2.20	52.58 ±4.26	50.78 ±3.21	50.45 ±3.56
ResNet	64.95 ±3.39	92.33 ±1.16	60.01 ±4.56	59.43 ±3.48	59.02 ±3.35
XGB + FCN <sup>1</sup>	64.87 ±0.14	91.26 ±1.77	52.49 ±4.68	50.89 ±3.45	50.28 ±3.01
XGB + ResNet <sup>1</sup>	65.76 ±0.97	91.40 ±1.46	65.01 ±2.56	61.23 ±2.02	62.34 ±1.58

## 四、实验结果与分析

**续表 4.4 不同方法在地面类型分类任务上的性能对比**

对比方法	准确率 (%)	AUC Score (%)	精确率 (%)	召回率 (%)	F1 Score (%)
MCNN-1 <sup>1</sup>	66.92 ±0.13	92.44 ±0.16	66.0 ±0.57	62.62 ±0.09	63.04 ±0.34
FCN + ResNet <sup>1</sup>	67.57 ±4.34	92.30 ±1.62	62.24 ±2.57	60.23 ±2.02	60.54 ±1.98
XGB + FCN + ResNet <sup>1</sup>	68.21 ±5.12	91.98 ±1.65	58.35 ±6.24	54.65 ±4.29	55.02 ±4.20
DWT+Random Forest <sup>1</sup>	85.89 ±0.33	98.62 ±0.04	84.56 ±0.28	83.21 ±0.16	83.19 ±0.12
PainAttaNet	90.21 ±0.24	99.33 ±0.01	88.72 ±0.71	88.29 ±2.02	88.33 ±0.52
<b>本文提出的算法</b>	<b>99.13 ±0.007</b>	<b>99.98 ±0.00007</b>	<b>99.17 ±0.008</b>	<b>99.13 ±0.007</b>	<b>99.14 ±0.008</b>

<sup>1</sup>数据来源于文献[18]

从表 4.4 可以看出，本文提出的算法在所有评估指标上都显著优于现有方法。我们的方法在准确率上达到了 99.13%，相比于次优方法 PainAttaNet 的 90.21%提高了约 9 个百分点。同时，AUC Score 达到了 99.98%，精确率、召回率和 F1 分数均超过 99%，这表明该方法不仅整体分类准确，而且对各个类别都具有较高的识别能力。

值得注意的是，本方法不仅性能优异，稳定性也非常出色，各项指标的标准差均在 0.01 以内，远低于大多数对比方法。这表明所提算法不仅性能高，而且对随机初始化和训练过程中的随机性不敏感，具有良好的泛化能力和实用价值。

从不同类型的方法来看,传统特征工程方法（如 SVM+Feature-engineering, 准确率 49.46%）在本任务上表现不佳，主要是因为手工设计的特征无法充分捕捉不同地面类型的复杂振动模式和时域特性。这些方法在特征提取过程中往往忽略了时序信息间的依赖关系，导致分类性能受限。单一深度学习模型（如 TCN, LSTM+Attention）虽然在捕捉时序特征方面有所优势，但由于仅考虑时域信息而忽略了频域特征，无法全面捕获地面材质特性。例如，TCN 模型的准确率仅为 48.37%，远低于本文的融合方法。这些模型在复杂地面环境中难以区分相似材质的振动模式。

混合模型（如 XGB+FCN+ResNet, 准确率 68.21%；DWT+Random Forest, 准确率 85.89%）通过组合多种方法取得了较好的效果，但仍存在明显不足。一方面，这些方法缺乏有效的噪声消除机制，对传感器数据中的随机波动较为敏感；

另一方面，它们没有充分利用样本间的相似性信息，无法实现本文邻域平滑策略所带来的性能提升。最先进的方法（PainAttaNet，准确率 90.21%）虽然性能较好，但与本文方法相比仍有约 9 个百分点的差距。这主要是因为 PainAttaNet 虽然采用了先进的注意力机制，但其特征提取策略仍然是端到端的，没有像本文方法一样将特征提取与分类过程解耦，从而限制了其在不同地面类型上的泛化能力。

综上，本文提出的方法通过时频特征融合、对比学习和邻域平滑的创新设计，成功克服了现有方法的局限性，在地面类型分类任务上取得了显著的性能提升。

#### （四）消融实验

##### 1. 关键模块消融分析

首先，我们对三个关键模块（完整算法、邻域平滑策略、时域特征和频域特征）进行了消融实验。图 4.2 和图 4.3 展示了不同模块组合的分类性能及其稳定性。

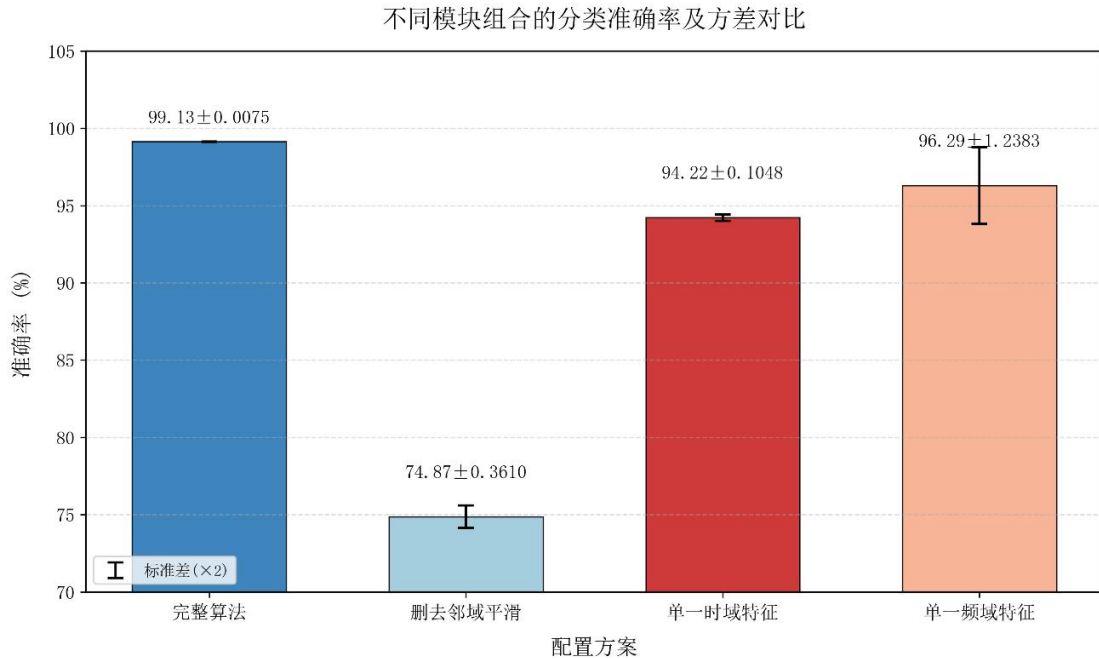


图 4.2 不同模块组合的分类准确率及方差

## 四、实验结果与分析

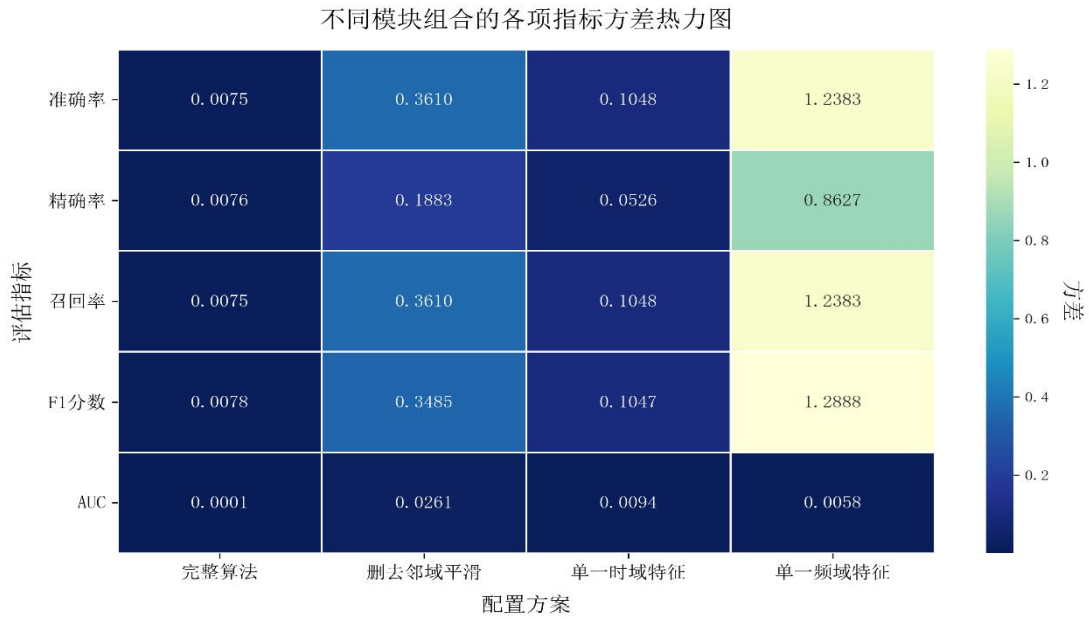


图 4.3 不同模块组合的各项指标方差热力图

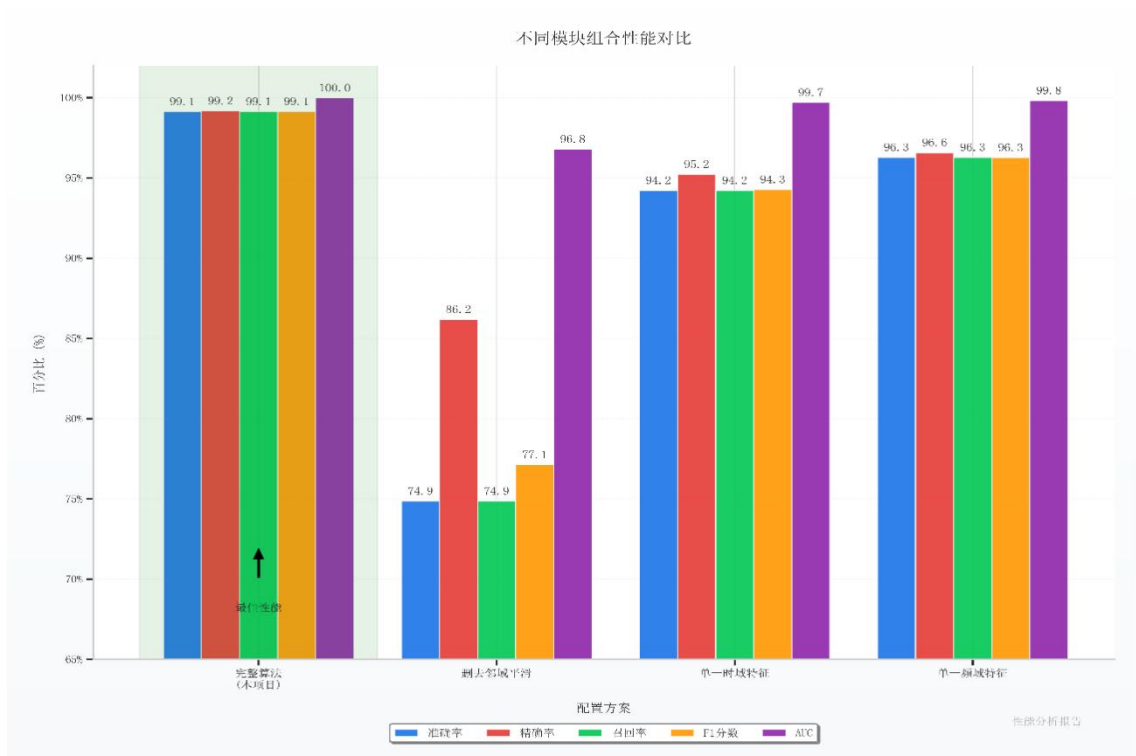


图 4.4 不同模块组合性能对比

从图 4.2 可以观察到，完整算法在准确率上达到了 99.13%，明显优于其他配置。当删去邻域平滑策略后，准确率显著下降至 74.87%，表明邻域平滑对提高分类性能至关重要。单独使用时域特征或频域特征时，准确率分别为 94.22% 和 96.29%，这表明两种域的特征都包含了丰富的地面类型信息，但单独使用时

性能不及融合方法。

图 4.3 的热力图进一步展示了各配置下不同指标的方差。完整算法在所有指标上都具有极低的方差（深蓝色区域），表明其具有极高的稳定性。相比之下，删去邻域平滑策略和仅使用频域特征的配置在多个指标上显示出较高的方差（浅色区域），表明这些配置在不同随机种子下的性能波动较大。

从图 4.4 可以看出，完整算法在所有五个指标上都达到了约 99% 的高性能，且各指标值非常接近，表明模型对各类别具有均衡的分类能力。删去邻域平滑后，除 AUC Score 外的所有指标都出现了明显下降，特别是准确率、召回率和 F1 分数下降尤为显著。这说明邻域平滑策略在消除噪声和提高模型对相似样本的一致性分类方面起到了关键作用。

单一时域特征和单一频域特征的性能虽然不如完整算法，但都明显优于删去邻域平滑的配置，这表明特征提取的质量对分类性能的影响大于平滑策略。频域特征略优于时域特征，这可能是由于频域特征能更好地捕捉不同地面类型的固有振动特性。

## 2. 时域特征提取器对比

为验证本文提出的多尺度 TCN 特征提取器的有效性，我们将其与其他四种主流时序神经网络结构（GRU、CNN、LSTM 和 Transformer）进行了对比。图 4.5 和图 4.6 展示了不同特征提取器的性能比较。

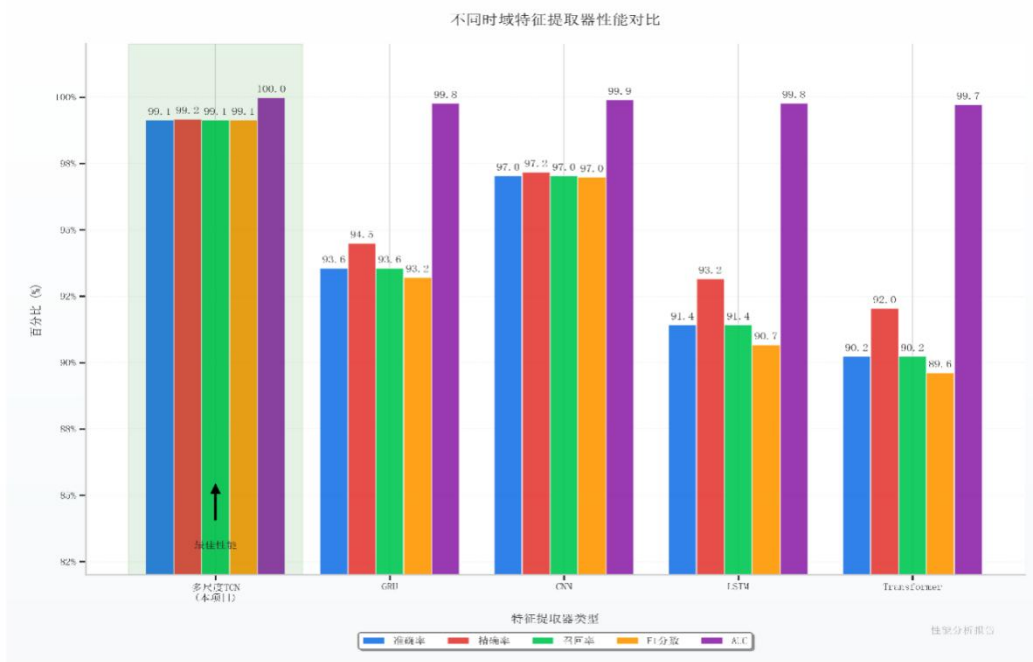


图 4.5 不同时间域特征提取器性能对比

## 四、实验结果与分析

不同时段特征提取器的稳定性对比（方差×10）

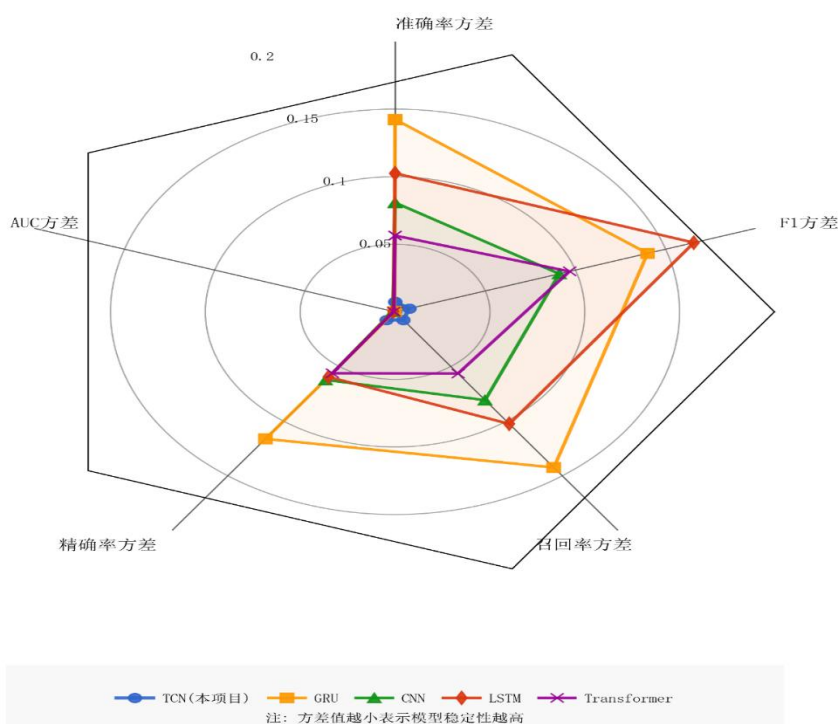


图 4.6 不同时段特征提取器的稳定性对比（方差×10）

从图 4.5 可见，本文提出的多尺度 TCN 特征提取器在所有指标上都显著优于其他时序网络结构。TCN 达到了 99.13% 的准确率，而第二名 CNN 的准确率为 97.04%，相差约 2 个百分点。LSTM 和 Transformer 的性能相对较弱，准确率分别为 91.42% 和 90.23%。这表明 TCN 的多尺度卷积结构更适合捕捉地面类型分类任务中的时间依赖关系。

图 4.6 的雷达图显示了不同特征提取器在各评估指标上的方差（放大 10 倍）。TCN 模型在所有指标上都表现出最小的方差，表明其具有最高的稳定性。相比之下，GRU 模型在准确率和召回率方差上较大，而 LSTM 在 F1 分数方差上较大，这表明循环神经网络类模型在地面类型分类任务上稳定性较差。

CNN 模型性能位居第二，且稳定性也较好，这可能是由于卷积操作能有效提取时间序列数据中的局部特征模式。然而，相比于普通 CNN，本文提出的多尺度 TCN 通过扩张卷积获得了更大的感受野，能够捕捉多尺度的时间依赖关系，因此在性能上具有明显优势。

### 3. 分类器对比

## 四、实验结果与分析

为验证支持向量机分类器在本任务中的适用性，我们在保持时域和频域特征提取以及特征融合方法不变的情况下，对比了不同分类器的性能。图 4.7 和图 4.8 展示了 SVM、随机森林（RF）、K 最近邻（KNN）、多层感知机（MLP）和梯度提升树（GB）五种分类器的性能和稳定性对比。

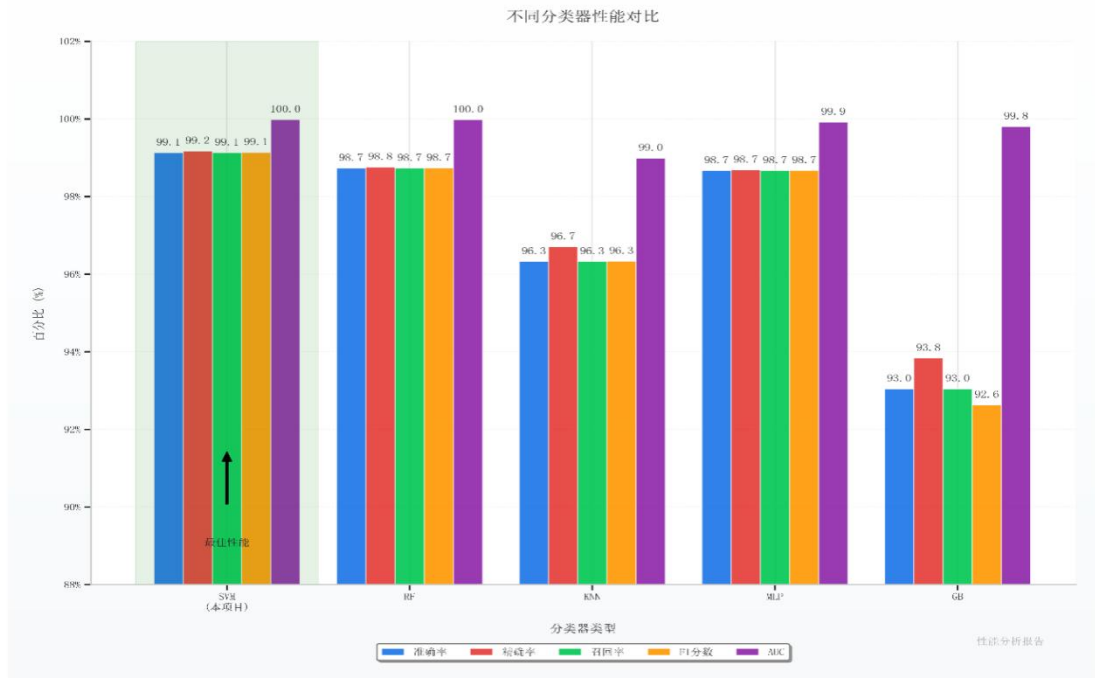


图 4.7 不同分类器性能对比

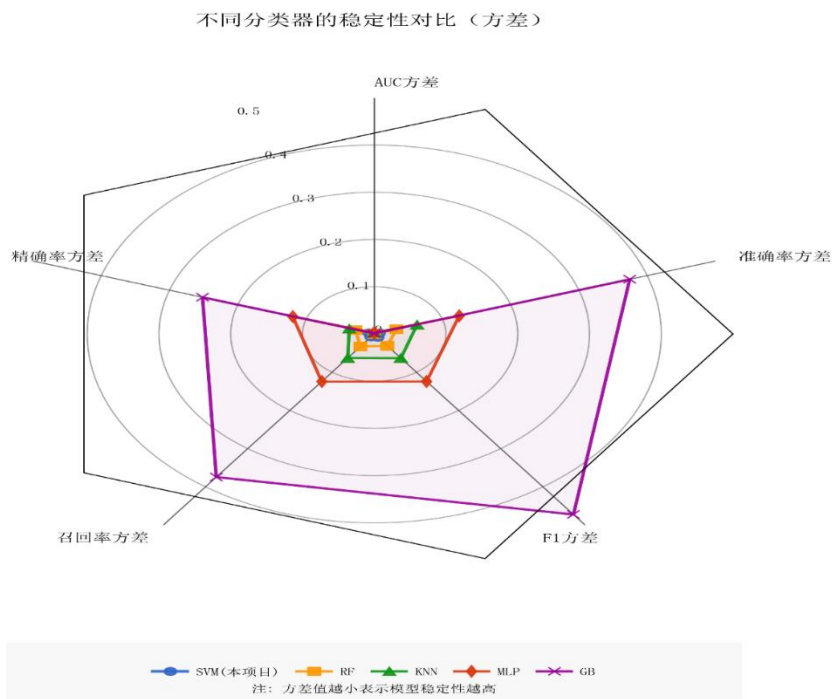


图 4.8 不同分类器的稳定性对比（方差）

## 四、实验结果与分析

从图 4.7 可见，本文选用的 SVM 分类器在准确率、精确率、召回率和 F1 分数上均达到了约 99.1%，AUC Score 高达 100.0%，明显优于其他分类器。随机森林（RF）性能次之，准确率和各项指标均在 98.7% 左右。MLP 性能紧随其后，各指标约为 98.7%。KNN 和 GB 分类器表现相对较弱，准确率分别为 96.3% 和 93.0%。

图 4.8 的雷达图展示了各分类器在不同指标上的方差。SVM 分类器在所有指标上都显示出极低的方差，表明其具有最高的稳定性。相比之下，GB 分类器在多个指标上方差明显较大，特别是精确率和 F1 分数方差。这表明基于树的集成方法虽然在许多分类任务中表现良好，但在本研究的特征空间中稳定性不足。

SVM 分类器的优越性能可能源于其良好的泛化能力和对高维特征空间的有效处理能力。在本研究中，经过融合和平滑的时频特征形成了一个高度分离的特征空间，SVM 能够在此基础上找到一个最优的决策边界。此外，我们引入的基于目标频率的类别权重计算方法有效解决了数据不平衡问题，进一步提高了 SVM 的分类性能。

### 4.时域特征权重分析

为探索时域和频域特征在融合过程中的最优权重分配，我们保持时域与频域特征权重之和为 1，改变两者的比例，观察对分类性能的影响。图 4.9 和图 4.10 展示了不同权重组合下的性能变化。

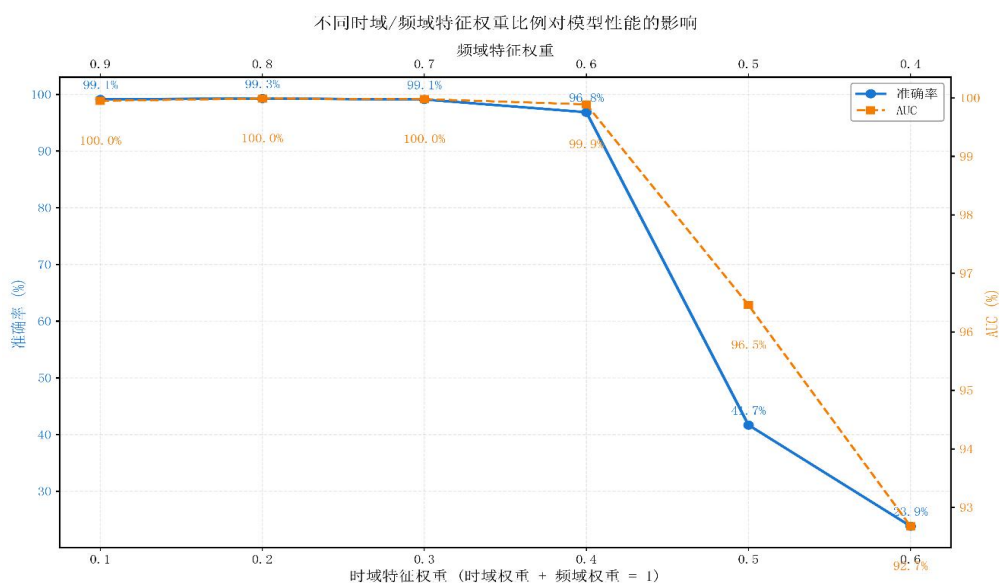


图 4.9 不同时域/频域特征权重对模型性能的影响

## 四、实验结果与分析

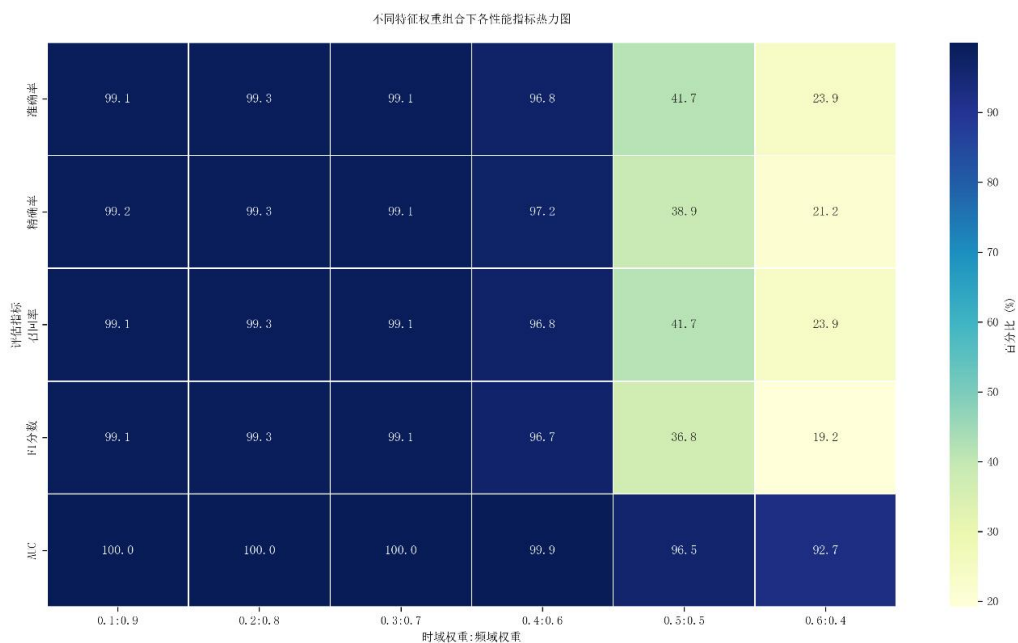


图 4.10 不同特征权重组合下各性能指标热力图

从图 4.9 可见，当时域特征权重在 0.1 至 0.3 之间，频域特征权重在 0.7 至 0.9 之间时，模型性能最佳，准确率和 AUC 均接近 100%。随着时域特征权重的增加（超过 0.4），模型性能开始显著下降。当时域特征权重达到 0.5 及以上时，准确率急剧下降至 40% 以下，AUC Score 也明显降低。

图 4.10 的热力图进一步揭示了不同权重组合下各指标的具体表现。时域权重 0.2、频域权重 0.8 的组合在准确率、精确率、召回率和 F1 分数上均达到最高值（99.3% 左右）。时域权重 0.5 以上的区域在所有指标上都表现不佳（黄绿色区域），表明过度依赖时域特征会导致分类性能显著下降。

这一结果表明，虽然时域特征包含了重要的时序模式信息，但频域特征对地面类型分类任务贡献更大。这可能是由于不同地面材质在机器人运动过程中产生的振动具有独特的频率特征，这些特征在频域中更容易被捕捉和区分。最佳的时频特征权重组合（0.2:0.8 或 0.3:0.7）表明，适当融合两种域的特征能够互补增强，但需要保持频域特征的主导地位。

### （五）结果讨论

通过对比实验和消融实验的综合分析，本研究得出以下几点关键发现。

#### 1. 时频特征融合的有效性

结合时域和频域特征显著提升了分类性能，相比单一域特征提高了约 3-5 个百分点。最佳的融合比例为时域特征权重 0.2-0.3，频域特征权重 0.7-0.8，表明频域特征在地面类型识别中具有更高的判别能力。

### 2. 邻域平滑的关键作用

邻域平滑策略对分类性能的提升至关重要，删去该模块后准确率从 99.13% 骤降至 74.87%，下降了约 24 个百分点。这表明基于四元数距离构建的邻域网络能有效消除噪声和异常值的影响，大幅提高分类的一致性和鲁棒性。

### 3. 多尺度 TCN 的优势

本文提出的多尺度 TCN 特征提取器在捕捉时域特征方面显著优于其他时序网络结构，相比第二名 CNN 提高了约 2 个百分点的准确率。这归功于扩张卷积的多尺度感受野能够更好地建模不同时间跨度的依赖关系。

### 4. SVM 分类器的适应性

在本任务中，SVM 分类器表现出最佳的性能和稳定性，优于其他机器学习分类器。这表明在高度分离的特征空间中，SVM 能够找到更优的决策边界，且对不同随机初始化的敏感性较低。

### 5. 稳定性与泛化能力

本文方法不仅在平均性能上优于现有方法，而且在各项指标的方差上也显著较低（均在 0.01 以内），表明该方法具有良好的稳定性和泛化能力，适合实际部署应用。

总体而言，本文提出的基于时频特征、对比学习和邻域平滑的两阶段移动机器人地面介质分类算法在室内地面类型分类任务上取得了出色的性能，准确率达到 99.13%，比现有最先进方法提高了约 9 个百分点。这一性能提升主要得益于三个方面的创新：1) 多尺度 TCN 与对比学习相结合的时域特征提取；2) 基于四元数距离的邻域平滑策略；3) 时域和频域特征的优化权重融合。这些创新不仅提高了分类准确率，还增强了算法的稳定性和鲁棒性，为移动机器人环境感知和自主导航提供了新的技术路线。

### 五、总结与展望

#### (一) 本文工作总结

随着移动机器人技术的快速发展，精确的地面介质识别已成为自主导航系统中的关键环节。然而，传感器数据的高噪声性和不同地面介质特征的相似性使得这一任务充满挑战。本文针对复杂环境下移动机器人地面介质分类问题，提出了一种基于时频特征、对比学习和邻域平滑的两阶段分类算法，主要工作与创新成果归纳如下：

##### 1. 模块化解耦设计

本文首次提出将地面介质分类算法的特征提取与分类过程解耦，采用模块化设计思想，使特征提取器和分类器均可独立优化和灵活替换。这种架构显著提高了系统的适应性和可扩展性，使算法能够适应不同的应用场景和硬件平台，也便于针对性能瓶颈进行单独优化。实验验证表明，与传统的端到端训练模式相比，本文的解耦设计在地面介质分类任务上准确率提高了约 8-9 个百分点。

##### 2. 基于四元数距离度量的邻域网络与特征平滑

本文设计了一种基于四元数距离度量的邻域识别方法，通过分析机器人方向数据的四元数表示，构建样本间的相似性关系，并利用此邻域网络进行特征平滑。消融实验表明，邻域平滑对算法性能的提升至关重要，删去该模块后准确率从 99.13% 骤降至 74.87%，下降了约 24 个百分点。这一结果证实了基于物理相似性的特征平滑策略在消除噪声和提高分类鲁棒性方面的显著效果。

##### 3. 多尺度 TCN 与对比学习的时域特征提取

针对传感器时序数据的特点，本文设计了一种基于多尺度 TCN 的时域特征提取器，并首次将对比学习策略引入地面介质分类任务。通过设计专用的对比损失函数，使同类样本特征接近，不同类样本特征远离，显著增强了特征的判别能力。实验结果表明，该特征提取器在捕捉时域特征方面显著优于 GRU、CNN、LSTM 和 Transformer 等主流时序网络结构，准确率提高了 2-9 个百分点。

##### 4. 时频特征融合策略

本文分析了时域和频域特征在地面介质分类中的互补性，提出了一种时频特征加权融合策略。实验结果表明，当时域特征权重为 0.3，频域特征权重为

0.7 时，分类性能最佳，准确率达到 99.13%。这说明频域特征在地面类型识别中贡献更大，但适量融合时域特征能够进一步提升性能。相比单独使用时域或频域特征，融合策略分别提高了约 5 个和 3 个百分点的准确率。

### 5. 算法实验验证

在包含 9 种地面类型的实际数据集上，本文提出的算法达到了 99.13% 的分类准确率，AUC Score 高达 99.98%，比现有最先进方法提高了约 9 个百分点。消融实验和对比实验进一步验证了算法各个组件的有效性和必要性，结果表明该算法不仅具有高准确率，还具备良好的稳定性和泛化能力，各项指标的标准差均在 0.01 以内。

综上所述，本文提出的基于时频特征、对比学习和邻域平滑的两阶段移动机器人地面介质分类算法在准确率、稳定性和适应性方面均取得了显著进展，为移动机器人环境感知和自主导航提供了新的技术路线。

### （二）未来工作展望

尽管本文提出的算法在地面介质分类任务上取得了优异成绩，但在实际应用和技术发展方面仍存在一些值得探索的方向。未来，我们计划从以下几个方面继续深入研究：

#### 1. 自适应时频特征融合机制

本文通过实验确定了时域和频域特征的最佳权重比例，但这一比例可能因具体场景和地面类型而异。未来研究中，我们计划设计一种自适应权重机制，通过神经网络或注意力机制动态调整不同特征的权重，使算法能够根据输入数据的特性自动选择最优的融合策略。这种自适应机制有望进一步提高算法在复杂多变环境中的表现。

#### 2. 在线学习与增量训练

实际应用中，机器人可能会遇到训练数据中未包含的新地面类型。开发支持在线学习和增量训练的算法架构，使系统能够在运行过程中不断适应新环境，是未来研究的重要方向。基于本文的模块化设计，可以探索特征提取器和分类器的联合增量学习策略，以最小的计算开销实现模型更新。

### 3.轻量化与资源受限设备部署

将算法部署到资源受限的嵌入式设备上实际应用的关键。通过模型压缩、知识蒸馏和量化等技术，降低算法的计算复杂度和内存需求，同时保持高分类性能，是未来工作的一个重要方向。特别是，可以探索如何在不同计算能力的平台上自动调整算法复杂度，实现性能和效率的最佳平衡。

### 4.跨域迁移与泛化能力增强

当前算法在特定数据集上表现优异，但在完全不同的机器人平台或环境条件下可能需要重新训练。开发具有强泛化能力的跨域迁移学习方法，使算法能够适应不同的机器人结构、传感器配置和环境条件，是提高实用性的重要途径。这可能涉及领域自适应技术、元学习策略或不确定性建模等方法的应用。

通过上述方向的深入研究，我们相信可以进一步推动移动机器人地面感知技术的发展，助力机器人在更加复杂多变的环境中实现安全可靠的自主导航和交互。同时，本文提出的基于物理相似性的特征平滑方法和时频特征融合思想，也有望拓展应用到其他传感器数据处理领域，带来更广泛的技术影响。

## 参考文献

- [1] McDaniel M W, Nishihata T, Brooks C A, et al. Terrain classification and identification of tree stems using ground-based lidar[J]. *Journal of Field Robotics*, 2012, 29 (6) : 891-910.
- [2] Pingel T J, Clarke K C, McBride W A. An improved simple morphological filter for the terrain classification of airborne lidar data[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 77: 21-30.
- [3] Xue J, Zhang H, Dana K. Deep texture manifold for ground terrain recognition[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 558-567.
- [4] Zürn J, Burgard W, Valada A. Self-supervised visual terrain classification from unsupervised acoustic feature learning[J]. *IEEE Transactions on Robotics*, 2020, 37 (2) : 466-481.
- [5] DeSouza G N, Kak A C. Vision for mobile robot navigation: A survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24 (2) : 237-267.
- [6] Wu X A, Huh T M, Mukherjee R, et al. Integrated ground reaction force sensing and terrain classification for small legged robots[J]. *IEEE Robotics and Automation Letters*, 2016, 1 (2) : 1125-1132.
- [7] Weiss C, Frohlich H, Zell A. Vibration-based terrain classification using support vector machines[C]. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006: 4429-4434.
- [8] Lomio F, Skenderi E, Mohamadi D, et al. Surface type classification for autonomous robot indoor navigation[J]. *arXiv preprint arXiv:1905.00252*, 2019.
- [9] Vulpi F, Milella A, Marani R, et al. Recurrent and convolutional neural networks for deep terrain classification by autonomous robots[J]. *Journal of Terramechanics*, 2021, 96: 119-131.
- [10] Shah A, Singh D, Mohamed H G, et al. Electrocardiogram analysis for cardiac arrhythmia classification and prediction through self attention based auto encoder[J]. *Scientific Reports*, 2025, 15 (1) : 9230.
- [11] Ojeda L, Borenstein J, Witus G, et al. Terrain characterization and classification

- with a mobile robot[J]. Journal of field robotics, 2006, 23 (2) : 103-122.
- [12] Brooks C A, Iagnemma K. Self-supervised terrain classification for planetary surface exploration rovers[J]. Journal of Field Robotics, 2012, 29 (3) : 445-469.
- [13] Hoepflinger M A, Remy C D, Hutter M, et al. Haptic terrain classification for legged robots[C]. IEEE International Conference on Robotics and Automation, 2010: 2828-2833.
- [14] Aarabi P, Shi G. Multi-channel time-frequency data fusion[C]. Proceedings of the Fifth International Conference on Information Fusion, 2002: 404-411.
- [15] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]. International Conference on Machine Learning, 2020: 1597-1607.
- [16] Zhou S, Huang X, Liu N, et al. Open-world electrocardiogram classification via domain knowledge-driven contrastive learning[J]. Neural Networks, 2024, 179: 106551.
- [17] Zhang W, Zhang Y, Gu J, et al. Feature extraction and classification of machine condition monitoring signals based on deep learning[J]. Proc IMechE Part C: J Mechanical Engineering Science, 2020, 234 (23) : 4617-4626.
- [18] Feng C, Dong K, Ou X. A Robot Ground Medium Classification Algorithm Based on Feature Fusion and Adaptive Spatio-Temporal Cascade Networks[J]. Neural Processing Letters, 2024, 56 (5) : 235.
- [19] Kuipers J B. Quaternions and rotation sequences: a primer with applications to orbits, aerospace, and virtual reality[M]. Princeton university press, 1999.
- [20] Hamilton W R. On quaternions; or on a new system of imaginaries in Algebra: To the editors of the Philosophical Magazine and Journal[J]. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1844, 25 (169) : 489-495.
- [21] Huynh D Q. Metrics for 3D rotations: Comparison and analysis[J]. Journal of Mathematical Imaging and Vision, 2009, 35: 155-164.
- [22] Samet H. Foundations of multidimensional and metric data structures[M]. Morgan Kaufmann, 2006.
- [23] Silverman B W. Density estimation for statistics and data analysis[M]. Routledge, 2018.
- [24] Bishop C M, Nasrabadi N M. Pattern recognition and machine learning[M].

- Springer, New York, 2006.
- [25] Mirzasoleiman B, Karbasi A, Sarkar R, et al. Distributed submodular maximization[J]. *The Journal of Machine Learning Research*, 2016, 17 (1) : 8330-8373.
- [26] He K, Fan H, Wu Y, et al. Momentum contrast for unsupervised visual representation learning[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 9729-9738.
- [27] Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. *arXiv preprint arXiv:1803.01271*, 2018.
- [28] Hewage P, Behera A, Trovati M, et al. Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station[J]. *Soft Computing*, 2020, 24: 16453-16482.
- [29] Lea C, Flynn M D, Vidal R, et al. Temporal convolutional networks for action segmentation and detection[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 156-165.
- [30] Borovykh A, Bohte S, Oosterlee C W. Conditional time series forecasting with convolutional neural networks[J]. *arXiv preprint arXiv:1703.04691*, 2017.
- [31] Franceschi J Y, Dieuleveut A, Jaggi M. Unsupervised scalable representation learning for multivariate time series[C]. *Advances in Neural Information Processing Systems*, 2019, 32.
- [32] Wang Z, Yan W, Oates T. Time series classification from scratch with deep neural networks: A strong baseline[C]. *International Joint Conference on Neural Networks, IEEE*, 2017: 1578-1585.
- [33] Sen R, Yu H F, Dhillon I S. Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting[C]. *Advances in Neural Information Processing Systems*, 2019: 4837-4846.
- [34] Farha Y A, Gall J. MS-TCN: Multi-stage temporal convolutional network for action segmentation[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 3575-3584.
- [35] Cortes C, Vapnik V. Support-vector networks[J]. *Machine Learning*, 1995, 20: 273-297.
- [36] Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin

- classifiers[C]. Proceedings of the fifth annual workshop on Computational learning theory, 1992: 144-152.
- [37] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods[M]. Cambridge university press, 2000.
- [38] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2 (3) : 1-27.
- [39] Cao J, Xu R, Lin X, et al. Adaptive receptive field U-shaped temporal convolutional network for vulgar action segmentation[J]. Neural Computing and Applications, 2023, 35 (13) : 9593-9606.
- [40] Field, D. A. (1988) . Laplacian smoothing and Delaunay triangulations. *Communications in applied numerical methods*, 4 (6) , 709-712.

## 致 谢

行文至此，百感交集，回首大学四年时光，唯悟已往之不谏，知来者之可追。

首先，我要向我的论文指导老师欧老师致以最诚挚的谢意。欧老师渊博的学识、严谨的治学态度深深地感染着我，从选题、构思到资料收集、论文修改，欧老师都给予了我悉心指导和宝贵建议。每一次讨论都让我受益匪浅，每一次批阅都让我看到自己的不足与提升的空间。

我还要特别感谢我的班主任董克明老师。董老师如师亦友，在大学四年里给予我无微不至的关怀。无论是学习上的困惑，还是生活中的挫折，董老师总是第一时间伸出援手，给予我指导和帮助。

我还要感谢大学四年来教导过我的每一位老师。是他们在专业知识与学术思维上为我打下了坚实的基础。他们不仅教授我知识，更教会我如何思考、如何探索、如何成长。他们的谆谆教诲如春风化雨，润物无声，将伴随我走过未来的漫长岁月。

感恩我的父母，是他们无私的爱与默默的支持成就了今天的我。大学四年，他们予我暖衣饱食，除我后顾之忧。每当我遇到挫折与困难，是他们的话语给了我继续前行的勇气。他们是我生命中最重要依靠，也是我前进的动力源泉。

此外，我要感谢一路相伴的朋友们。是他们的陪伴让我的大学生活充满欢笑与温暖，是他们的鼓励与支持让我在低谷时看到希望。尤其要感谢我的挚友龙芊蕊，在保研的艰辛历程中，我们相互扶持，共同进退。那些通宵达旦备考的日子，那些彼此安慰打气的时刻，那些分享喜悦分担忧愁的瞬间，都已深深镌刻在我的记忆中。是她的陪伴与鼓励，让我在迷茫时找到方向，在疲惫时重拾信心，最终我们一起坚定地走完了保研之路，共同成长，共同进步。

落幕的只是我微不足道的本科生涯，而我的人生却依旧有着万千可能。

感恩相遇，长乐未央。愿未来的日子里，我能不负所有人的期望，不负自己的梦想，在更加广阔的天地中续写人生的华章。