

考虑多样性和可辨别性表示学习的低资源可泛化活动识别模型

摘要

人类活动识别 (Human Activity Recognition, HAR) 在普适计算中发挥着很重要的作用。知识的迁移在 HAR 的场景中是常见的一项技术, 但获取足够的训练数据需要很高的时间和经济成本, 因此数据资源的匮乏是目前实现可泛化的 HAR 模型的主要瓶颈之一。另外, 不同的人群由于行为习惯、年龄、体型等差异会导致数据在分布上会存在偏移。低资源和分布偏移的问题成为目前 HAR 的迁移学习中最主要的两大问题。针对这两大挑战, 本文基于域泛化的方法设计了一个应对低资源的可泛化 HAR 模型, 模型通过数据增强、对比学习等技术将数据的多样性和特征的可辨别性考虑进来, 使得模型在学习到数据多样性的同时能够提取到域无关的特征。实验表明, 本文的模型对比现有方法具有良好的性能和泛化能力。

关键词: 人类活动识别; 域泛化; 低资源

1 引言

人类活动识别 (HAR) 在许多普适计算的应用中都发挥着不可或缺的作用。常见的应用场景有睡眠监测、手势识别、驾驶过程的疲劳监测等等。这些应用要求研究者设计通用且准确的 HAR 算法, 从而使得在源用户数据上训练的模型能够很好地迁移到目标用户上, 实现良好的预测效果。然而, 由于 HAR 需要的标签数据需要在特定的实验环境下人工采集, 这导致标签数据的收集十分耗时耗力。而且由于我们无法全天候采集人体的所有活动数据, 因此数据的多样性是不足的。这进一步造成了模型难以训练出良好的泛化能力。另外, 不同实验人群由于年龄、体型、行为习惯, 甚至不同的设备佩戴位置的差异, 他们在进行相同类型的活动时, 数据的分布规律可能是不同的。这种数据的分布偏移也对泛化模型提出了挑战。目前许多工作围绕上述两大挑战展开了很多研究, 也取得了一定的成果。但是现有方法大都以域适应的思想构建模型, 本文将从现实出发, 以更加实际的角度, 基于域泛化的方法应对上述两个挑战。

2 相关工作

2.1 人类活动识别

人类活动识别主要是对人体日常生活中发生的不同活动进行分类。根据采用的数据类型，HAR 大体可以分为基于传感器和基于视觉两大类。本课题专注于基于传感器数据的 HAR。基于传感器的 HAR 通过环境部署的传感器或可穿戴传感器捕获活动数据。目前可穿戴式的设备广泛部署惯性测量单元（Inertial Measurement Unit, IMU），使得记录人体日常活动数据变得更加方便 [1]。

2.2 域泛化

域泛化（Domain Generalization, DG）是迁移学习当中的一项技术，旨在从一个或多个具有不同概率分布的源域中学习一个健壮且可泛化的模型，并使其在未知的目标域获得最小的误差。与域适应（Domain Adaptation, DA）不同，域泛化假设在训练阶段无法接触到目标域的数据，直接将训练好的模型应用到目标域上，没有域适应当中在目标域上再训练或微调的阶段 [2,8]。

现有的 DG 方法可以分为三大类，分别是数据操作 [4]、表示学习 [3] 以及学习策略 [5]。尽管在视觉领域 DG 方法得到很广泛的应用，但是在 HAR 任务中，DG 的工作还比较少。目前第一个使用 DG 来解决 HAR 问题的工作是通过解开领域不可知和领域特定的特征提出的 [6]。然而，现有的方法仍然很少尝试解决可推广的低资源 HAR 问题。

3 本文方法

3.1 本文方法概述

本文的模型框架如图 1 所示。主要分为模型的训练和推理两个阶段。在训练阶段主要分为四个步骤，分别是多样性生成、特征提取、多样性保留和可辨别性增强，以及最后的分类。最终目标是训练出一个特征提取器和活动分类器，并在推理阶段应用于目标域。

3.2 训练阶段

多样性生成。这里采用 7 种不同的数据增强方法，分别是 jittering、scaling、magnitude warping、time warping、rotation、permutation 和 random sampling。增强后的数据与原始数据构成 8 个不同的类别，在分类步骤将采用一个自监督辅助分类器来辨别数据来自哪一类，从而提升特征提取器对数据关键特征的提取能力。

特征提取。特征提取器采用两层二维卷积网络以全连接的方式搭建，提取数据的特征，同时根据其他部分的损失更新自身参数。

多样性保留 & 可辨别性增强。由于分类目标的影响中，在训练过程中相同类别的表示空间可能会靠近缩小，这会导致数据增强得到的数据多样性减少。因此，本文设计了一个二元分类器辨别原始数据和增强数据，通过最小化辨别器的误差，原始数据与增强数据的距离可以增大，从而区别原始数据和增强数据。另外，增强表示的语义区分对实现准确的活动识别

非常重要。本文通过使用监督对比学习的方法，扩大不同类之间的距离，并减少相同类内的距离来增大类之间的区分性，从而实现更准确的活动识别。

分类。该步骤最主要是训练活动分类器，即对输入数据的活动标签进行预测。

3.3 推理阶段

在推理阶段，我们将使用前面训练好的特征提取器和活动分类器，将目标域的数据输入特征提取器，对活动标签进行预测。

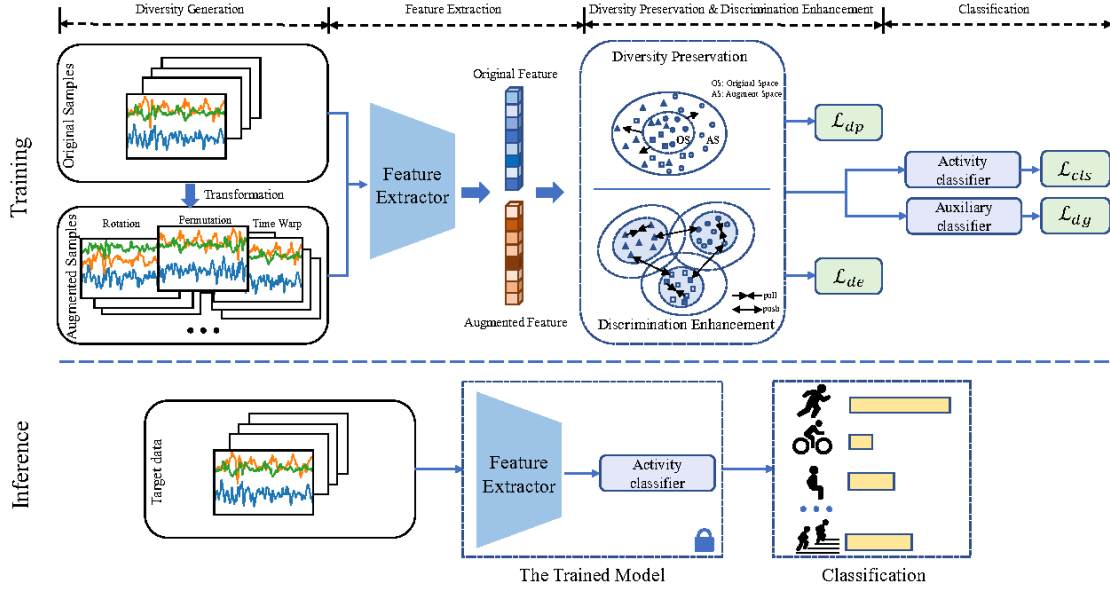


图 1. 模型框架图

4 复现细节

4.1 与已有开源代码对比

本论文提供了源代码，本课题的代码是参考源代码重新编写而成。本课题对项目的复现在源代码的基础上，增加了一个 HAR 研究常用的数据集 HHAR [7]，并且重新设计了实验部分对工作进行性能测试。复现的流程如图 2 所示。

4.2 实验环境搭建

本课题使用的编程语言是 Python 3.9，并且使用深度学习框架 Pytorch 2.1.0 搭建模型，在 Windows10 的操作系统，PyCharm 平台进行编程。机器的 CPU 和 GPU 配置分别是：Intel(R) Core(TM) i7-10700K CPU @ 3.80GHz(16 CPUs)，NVIDIA GeForce RTX 3090。

4.3 创新点

1. 实验部分新增 HHAR 数据集，测试模型泛化能力。
2. 重新设计实验，对每个数据集在不同百分比的训练数据下进行各项迁移任务的测试。

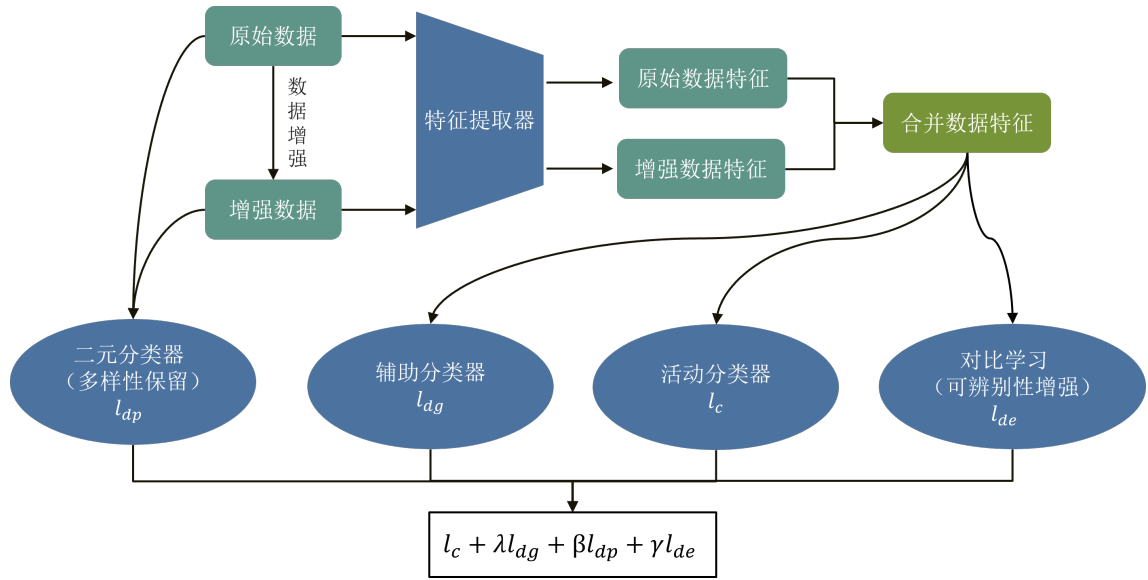


图 2. 模型框架图

5 实验结果分析

实验按照数据集大小将各个数据集划分成 3 到 5 个不等的迁移任务，对每个任务分别在有 20% 和 100% 的训练数据下测试了模型的预测准确度。实验结果用表格和柱状图的形式展示，分别是表 1 和图 3。相比原文只在 20% 的训练数据下进行测试，本课题设计的实验可以更直观的看到模型对低资源情况的提升。我们可以看到，对每个数据集，20% 的训练数据下的模型准确度相比 100% 的准确度并没有很大差异，说明了模型具有良好的泛化能力。

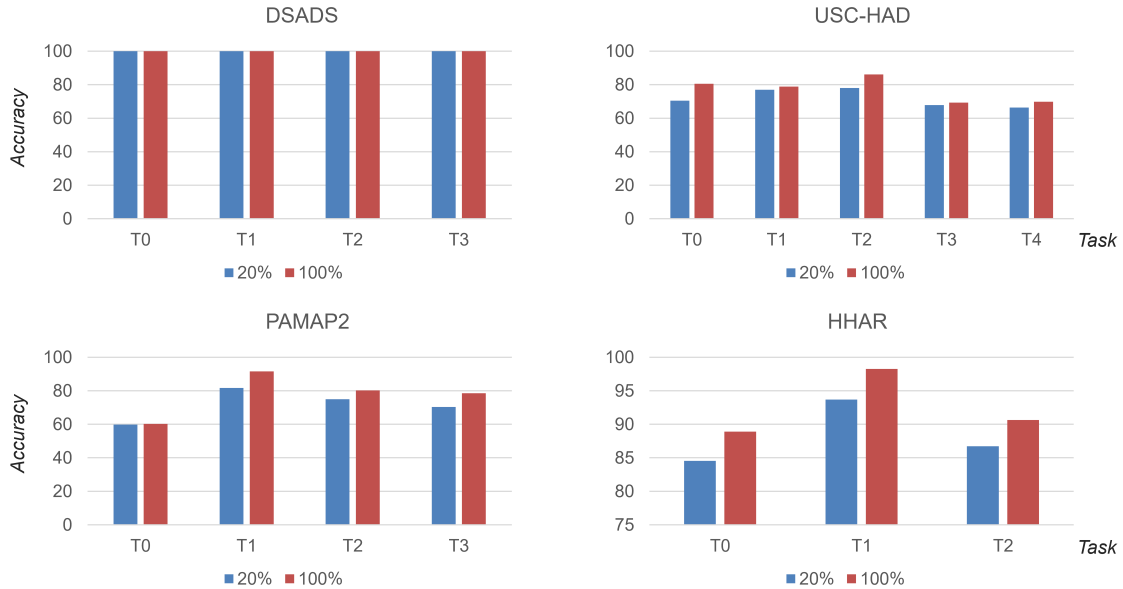


图 3. 模型在不同百分比训练数据下的准确度

表 1. 模型在不同百分比训练数据下的准确度

数据集	迁移任务	训练数据百分比	准确度
DSADS	T0	20%	99.10
		100%	100.0
	T1	20%	98.87
		100%	100.0
	T2	20%	98.30
		100%	100.0
	T3	20%	97.26
		100%	100.0
PAMAP2	T0	20%	59.85
		100%	60.22
	T1	20%	81.62
		100%	91.54
	T2	20%	74.91
		100%	80.21
	T3	20%	70.29
		100%	78.46
USC-HAD	T0	20%	70.50
		100%	80.58
	T1	20%	76.96
		100%	78.92
	T2	20%	78.03
		100%	86.11
	T3	20%	67.87
		100%	69.32
HHAR	T0	20%	84.53
		100%	88.89
	T1	20%	93.68
		100%	98.26
	T2	20%	86.71
		100%	90.63

6 总结与展望

本课题复现了发表在 SIGKDD 2023 的论文《Generalizable Low-Resource Activity Recognition with Diverse and Discriminative Representation Learning》，旨在使用域泛化的方法，考

虑数据多样性和可辨别性的同时,实现应对低资源的可泛化 HAR 模型。在未来的工作中,我们可以进一步考虑特征提取器的神经网络设计,采用新颖有效的网络结构,提升特征提取的能力。另外,可以在分类步骤根据需求添加辅助分类器,从而帮助数据的特征提取。

参考文献

- [1] Yash Jain, Chi Ian Tang, Chulhong Min, Fahim Kawsar, and Akhil Mathur. Collossl: Collaborative self-supervised learning for human activity recognition, 2022.
- [2] Wang Lu, Yiqiang Chen, Jindong Wang, and Xin Qin. Cross-domain activity recognition via substructural optimal transport. *Neurocomputing*, 454:65–75, September 2021.
- [3] Wang Lu, Jindong Wang, and Yiqiang Chen. Local and global alignments for generalizable sensor-based human activity recognition. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, page 3833–3837, 2022.
- [4] Wang Lu, Jindong Wang, Yiqiang Chen, Sinno Jialin Pan, Chunyu Hu, and Xin Qin. Semantic-discriminative mixup for generalizable sensor-based cross-domain activity recognition, 2022.
- [5] Wang Lu, Jindong Wang, Xinwei Sun, Yiqiang Chen, and Xing Xie. Out-of-distribution representation learning for time series classification, 2023.
- [6] Hangwei Qian, Sinno Jialin Pan, and Chunyan Miao. Latent independent excitation for generalizable sensor-based cross-person activity recognition. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021*, page 11921–11929, 2021.
- [7] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems, SenSys '15*, page 127–140, 2015.
- [8] Garrett Wilson, Janardhan Rao Doppa, and Diane J. Cook. Multi-source deep domain adaptation with weak supervision for time-series sensor data, 2020.