研究生开题报告书

|  |
| --- |
| 1. 选题依据：（论文的研究意义、国内外研究现状分析）   [1] ROY P K, OM H. Suspicious and violent activity detection of humans using HOG features and SVM classifier in surveillance videos[J]. Advances in Soft Computing and Machine Learning in Image Processing, 2018: 277–294.  [2] EHATISHAM-UL-HAQ M, JAVED A, AZAM M A, et al. Robust human activity recognition using multimodal feature-level fusion[J]. IEEE Access, 2019, 7: 60736–60751.  [3] RAFFERTY J, NUGENT C D, LIU J, et al. From activity recognition to intention recognition for assisted living within smart homes[J]. IEEE Transactions on Human-Machine Systems, 2017, 47(3): 368–379.  [4] FUKUSHIMA K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36(4): 193–202.  [5] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735–1780.  [6] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.  [7] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. Journal of machine learning research, 2016, 17(59): 1–35.  人类行为识别介绍：随着科学技术的发展，人类活动识别（Human Activity Recognition，HAR）在许多领域引起了比以往更多的关注，例如安全[1]、医疗保健[2]和智能家居[3]。一般来说，HAR 可以分为基于视觉的 HAR 和基于传感器的 HAR。前者通常依赖于摄像头，例如 RGB 摄像头或深度摄像头，而后者依赖于可穿戴传感器，例如加速度计、陀螺仪和磁力计。本文将讨论基于传感器的 HAR。基于传感器的HAR将从若干位受试者身上佩戴的传感器，并在固定时间内做出指定活动，从而取得每一位受试者的多个活动的传感器数据。HAR算法通过有监督学习，将通过包含传感器数据的固定长度的滑动窗口，判断该滑动窗口所属的活动类别。 以下提到的所有HAR均表示基于传感器的HAR。  目前有两种解决HAR的解决方案：   1. 一种解决方案假设数据满足独立同分布（Independent and Identically Distributed，I.I.D）。将来自所有受试者的数据混洗并划分出固定的训练集和测试集。这种角度下存在机器学习和深度学习两种方法。机器学习通过领域知识制作人工特征，并使用机器学习方法完成HAR，例如随机森林、XGBoost等，主要关注手工制作的特征，而这些特征大多依赖于领域专业知识；深度学习则通过神经网络学习数据的隐藏特征，不仅不需要领域知识，还能让HAR达到更高的准确率。卷积神经网络[4]（Convolutional Neural Network，CNN）和长短期记忆单元[5]（Long Short-Term Memory，LSTM）常被用于解决HAR，卷积神经网络用于学习多个传感器特征在同一个时间步内的空间特征，而长短期记忆单元用于学习同一个传感器特征在多个时间步内的时间特征，结合空间特征和时间特征提高HAR的准确率。不仅如此，自注意力机制（Self-Attention Mechanism）通过神经网络学习特征的内部相对重要性，来进一步提高深度学习特征的表示能力。   上述解决方案介绍了独立同分布的HAR。然而，不同人的活动数据可能是不同分布的，这也称为分布外（Out-of-distribution），又称领域偏移（Domain Shift），领域偏移将导致模型在HAR上的准确率显著降低。如图1所示，通过佩戴在右臂上的智能手表中的加速度计收集到四名受试者的行走数据，橙色、绿色、蓝色分别表示加速度计的x、y、z轴的分量。其中，穿黄色衣服的人在z轴加速度上波动较大，而穿橘色衣服的人在z轴加速度上几乎没有波动，年龄、身高、体重、行为习惯等的差异造成均有可能导致如图1所示的领域偏移。    图1   1. 因此，另一种解决方案假设数据分不满足独立同分布（Independent and Identically Distributed，I.I.D），从而侧重于HAR中的解决领域偏移。其中，每名受试者的数据被视为一个领域，领域与领域的数据分布相互独立且不同。这种解决方案使用迁移学习解决领域偏移问题，迁移学习的目标是在数据分布变化时最小化泛化误差。例如，在计算机视觉的图像分类中，真实生活背景和卡通纯色背景下的键盘图片都应指向同一个标签“键盘”，迁移学习将帮助模型从真实生活背景的图片中学习到键盘的样子，并在背景发生变化时依然能识别出键盘。在HAR 中，迁移学习主要分为领域适应 (Domain Adaption，DA) 和领域泛化 (Domain，Generalization，DG) 。领域适应的目的是将测试数据视为未标记数据，并通过某种可学习的方式设法减少训练标记数据和测试未标记数据之间的差异。受到生成对抗网络[6] (GAN) 的启发，领域对抗神经网络[7](DANN)成为传统半监督学习的一个很好的例子，DANN中的生成器和鉴别器协同工作，在隐藏特征空间缩小领域差异。然而，领域适应中使用测试数据存在数据泄露的隐患，未能遵循“只需训练一次即可随处测试”的原则。因此，在不使用测试数据的前提下，领域泛化逐渐成为 HAR 中的热门研究方向。领域泛化的目标是在训练期间没有测试数据可用的情况下实现领域适应的相同目标。领域泛化更重视事物的本质，回到图1，领域泛化将试图提取每个人走路的周期性变化特征，并忽略每个人的步幅等个人习惯因素。     图2  论文的研究意义：本研究设计了一种用于HAR领域泛化的特征增强框架，图2展示了包含该框架在内的HAR领域泛化流程。该框架包含四个部分，分别是特征提取网络、元学习模块、特征解耦模块和分类网络。特征提取网络用于提取任何和人类行为相关的特征；元学习模块用于限制特征提取网络，使其提取有利于克服领域偏移的领域不变特征；特征解耦模块在此基础上，使特征提取网络不仅提取领域不变特征，还从中尽可能提取领域可变特征，从而将领域可变特征和领域不变特征解耦；分类网络通过领域不变特征得到数据对应的人类行为标签。本论文从数据样本层面上升到领域层面，用颗粒度更细的领域泛化的角度看待HAR，旨在减轻甚至克服因受试者行为习惯不同导致的数据分布不同，加强HAR模型在不同场景下的预测稳定性，扩展了HAR模型在训练完后的泛化能力，提高HAR预测的准确率。  国内外研究现状：   1. 基于传感器的人类行为识别：   基于传感器的人类活动识别的总体目标是将固定的滑动窗口分类为几个预定义的活动。滑动窗口最初由多个传感器（例如加速度计、陀螺仪等）获得。基于手工制作的特征，机器学习方法被用于处理基于传感器的 HAR 问题，例如决策树、支持向量机和隐马尔可夫模型。除了手工制作的特征外，深度神经网络学习到的特征在缺乏领域专业知识的情况下也能很好地解决 HAR 问题。用于 HAR 的 CNN 首先利用几个内核大小为 1x1的深度 CNN 来提取滑动窗口内时间戳上的潜在特征。在 CNN 层后面添加循环神经网络 LSTM，以进一步探索每个滑动窗口内所有时间戳之间的相对重要性。之后，CNN-RNN 的混合结构在 HAR 问题上大获成功。CNN-RNN 结构的多种变体被提出，例如多分支 CNN-LSTM。此外，随着注意力机制的发展，注意力的变体被广泛引入到混合结构中。在 CNN-LSTM 结构后面添加了一个自注意力模块，并通过内部比较来细化和聚焦学习到的特征。被提出来不仅能获得通道注意力，还能获得空间注意力，他们还在自己构建的数据集上对其进行了测试，结果证明 DanHAR 是有效的。三重注意力被提出来考虑通道、时间戳和特征之间的关系。此外，MTSDNet 被提出来通过学习系数来仅使用全连接层、傅里叶级数和幂级数来提取时间序列特征。   1. 迁移学习：   迁移学习的主流方法分为领域适应（Domain Adaption，DA）和领域泛化（Domain Generalization，DG）。迁移学习中的大多工作为图像分类设计，在基于传感器的HAR也存在一些适配的方法，下述文字将一并介绍迁移学习在图像分类中的原始方法和在HAR中的适配方法。  领域适应的目的是充分利用未标记的目标域中的知识以实现更好的迁移，而 DG 则试图仅从源域中学习更多知识并保持目标域不可见。基于传感器的HAR 的主流 DA 方法侧重于对抗性学习，将未标记的目标域迁移到与源域相似的域中，从而减少源域和目标域之间的差异。DANN试图通过神经网络减少源域和目标域之间的差异。CORAL通过源域和目标域之间的相关矩阵差异来对齐域。对抗学习自编码器通过将目标生成到可控域中来对齐源和目标。同样，【文献】提出了具有不同判别器判断条件的GAN，以可学习的方式进一步减少源和目标之间的差异。  基于传感器的HAR的主流DG方法主要集中在数据增强，特征对齐、元学习和特征解耦上。数据增强是通过不同方法增加数据样本的丰富程度，从而减少Mixup在不同样本之间进行线性插值，从而丰富了样本的多样性。语义混合方法为线性插值添加了活动语义含义，以排除无意义的生成样本。对于领域不变特征学习，GILE 使用了两个变分自动编码器，以使用独立的激励机制同时学习领域变异和领域不变特征。RSC将某些相关特征的梯度设置为零，以鼓励模型学习领域不变特征。元学习最初为图像分类和文本分类的少样本学习服务，其目的是让模型学会学习，根据模型的临时梯度下降在多份数据上的表现，从而决定正式梯度下降的方向。 模型无关的元学习 (Model Agnostic Meta-Learning，MAML) 是一个用于小样本学习和强化学习的元学习框架。MAML 有助于找到对新任务变化敏感的参数的良好初始化。最近几年来，基于 MAML 的方法已广泛应用于各种任务，例如领域泛化、对象跟踪 和帧插值等，此处仅介绍为领域泛化服务的元学习方法。此从 DG 的角度来看，MLDG是一种基于 MAML 的代表性方法，它能够通过模拟真实的训练测试域转移来训练具有良好泛化能力的模型到新领域。此外，Balaji 等人 在元学习管道中训练正则化器以实现良好的跨域泛化。  强制整个模型学习领域不变特征是有挑战性的，特征解耦在学习领域不变特征的基础上，允许某些部分特定于领域，本质上就是学习解耦表示。特征解耦存在基于分解的方法和基于生成模型构建的方法，两者都需要领域标签来实现，即模型需要知晓所有数据分别所属的领域。基于分解的方法将模型分解为两部分，一部分是领域特定的，另一部分是领域无关的。基于 SVM，Khosla 等人 [46] 将分类器分解为领域特定的偏差和领域无关的权重，并且只在处理看不见的域时保留后者。这种方法后来在 [37] 中扩展到神经网络。我们还可以设计领域特定模块，例如 [207] 中的方法，其中在最终特征向量上施加领域特定二进制掩码，以区分领域特定和领域不变成分。另一种解决方案是对模型的权重矩阵进行低秩分解，以识别更具泛化的共同特征 [208]。生成式建模一直是学习解缠表示的有力工具 [246]。在 [209] 中，利用变分自编码器 (VAE) 分别学习类、域和对象的三个独立潜在子空间。在 [210] 中，以对抗的方式学习两个独立的编码器，分别捕获身份和域信息，用于跨域人脸反欺骗。 |
| 二、研究方案设计：  1.研究目标、研究内容和拟解决的关键问题 。  2.拟采取的研究方法、技术路线、实施方案。  3.论文特色与创新点。  4.研究的预期成果与进度安排。  1.1研究目标  本研究设计了一种用于HAR领域泛化的特征增强框架，该框架包含四个部分，分别是特征提取网络、元学习模块、特征解耦模块和分类网络，用于提高领域泛化角度下HAR的预测准确率。  1.2 研究内容  1. 本研究围绕基于传感器的HAR（下均简称为HAR）中存在的领域偏移问题，传统任务假设数据是独立同分布的，而该问题假设数据不满足独立同分布。其中，每名受试者的数据被视为一个领域，领域与领域的数据分布相互独立且不同。领域偏移问题是一个现实问题，即现实生活中确实存在因行为习惯导致的人类行为差异。  2. 本研究的数据基础由两部分，一是多个热门公共HAR数据集，一是自建HAR数据集。本研究希望在此数据基础上得到以下几个目标：首先，使用本研究所设计的特征提取框架，训练的HAR模型的具有更优秀的预测指标，指标是在每个领域上评估得到多个指标，指标可以是一种或多种，在每个领域上的平均指标更高表明该框架更优秀；其次，特征提取框架中存在领域不变特征和领域可变特征，前者应被可视化，用于证明该框架能够提取数据集内同一种人类行为在不同领域之间存在的不变特征；最后，在模型预测指标和可视化结果的基础上，本框架应可被用于分析每个受试者的行为差异性，即数据集的中受试者之间是否存在行为差异的定性分析。  1.3 拟解决的关键问题  1. 特征提取网络应选取HAR领域最新的热门方法，保持本研究的时效性。  2. 元学习模块应合理结合现有的元学习方法和HAR领域知识，在保持方法与适配的情况下防止元学习不收敛。  3. 特征解耦模块必须合理设计损失函数，避免模型同时学到性质相同的特征（同时学到两种领域无关特征，说明特征解耦失败），防止模式崩坏。  4. 合理设计领域无关的可视化，从而足以证明所设计框架的有效性   * 1. 拟采取的研究方法   数据集：   1. 在公开的热门数据集上展开实验 2. 自建一个弱标注的数据集验证两种框架的有效性   框架设计方法：   1. 借鉴图像分类领域，汇总现有的元学习方法，汇总现有的为解决领域泛化的元学习方法，将合适的元学习方法适配在人类行为识别领域。 2. 调研特征解耦方面，近5年的文献，包括情感分类、图像分类、目标检测等，总结特征解耦的方法特点   实验设计方法：   1. 通过分类指标判断当前方法的可行性 2. 通过相关工作的分类指标比较当前方法的优越性 3. 通过分类指标和当前方法的消融实验，判断当前方法内各部分模块的必要性   1. 将近三年来  2.2 技术路线  2.2.1 数据集  公开数据集处理：将公开的可穿戴传感器数据集按领域划分，并划分为滑动窗口。  将公开的可穿戴传感器数据集中的若干个人类志愿者等效视为若干个领域，并假设任意领域之间的传感器数据联合分布均不同；随后，将每个领域的传感器数据进一步划分为若干个等长的滑动窗口。  具体的，对含有N个人类志愿者的数据集划分为个领域，对任意一个领域，均有个时间步的传感器数据和人类行为标签，和具有相同的时间步，即每条传感器数据均存在人类行为标签。滑动窗口记作和，分别表示第个领域的第个滑动窗口和滑动窗口对应的人类行为标签，其中，表示由第个领域划分出滑动窗口的总数，即, 为所有领域的滑动窗口包含的时间步数，为每个滑动窗口具有的传感器特征数；滑动窗口和滑动窗口对应的人类行为标签是人类行为识别输入至模型的基本数据单位。  本步骤中，滑动窗口的向量化详细过程为：首先，提取滑动窗口均值和标准差（即滑动窗口的统计特征），并标准化滑动窗口。  自建数据集收集方法：使用IOS的软件phyphox收集并导出数据集，该软件可调用IOS手机内部的加速度计和陀螺仪，收集加速度计和陀螺仪在x、y、z方向上的数据，因此在该自建数据集中。数据集预计收录5种简单活动，包括走路、跑步、静坐、上楼和下楼等，  2.3 实施方案  3 论文特色与创新点  本研究在方法上首次将元学习与人类行为识别的领域偏移问题结合，从深度学习框架的角度改变模型学习流程，提取领域无关的特征，从而提高领域偏移中模型在人类行为识别上的分类指标。另外，本研究首次将特征解耦与元学习结合，从领域无关特征中分离了领域相关特征，进一步增强了领域无关特征的表示。此外，本研究还自建了一个包含简单活动的数据集，从数据角度验证本研究的有效性。 |
| 三、论文大纲（至少二级提纲）： |
| 四、研究的前期基础与可行性分析  文献阅读基础  技术基础  技术可行性  环境可行性  时间可行性 |
| 五、中英文参考文献（至少40篇以上）  六、附件  张老师：   1. 标题题目过于模糊，无法一眼获取研究的大致内容。   回复：   1. 为什么没有在行为语义方面深入研究？   回复：   1. 收集数据集不应该只局限于自己使用。   回复：  丁老师：   1. PPT中所示元学习是否实际上是增量学习 |