



(12) 发明专利申请

(10) 申请公布号 CN 118585901 A

(43) 申请公布日 2024. 09. 03

(21) 申请号 202411074110.2

(22) 申请日 2024.08.07

(71) 申请人 深圳市大数据研究院

地址 518172 广东省深圳市龙岗区龙城街道龙翔大道2001号道远楼225室

(72) 发明人 赵子健 朱光旭 陈琪美 韩凯峰

(74) 专利代理机构 广州嘉权专利商标事务有限公司 44205

专利代理师 周翀

(51) Int. Cl.

G06F 18/2413 (2023.01)

G06F 18/214 (2023.01)

G06F 18/21 (2023.01)

权利要求书3页 说明书20页 附图4页

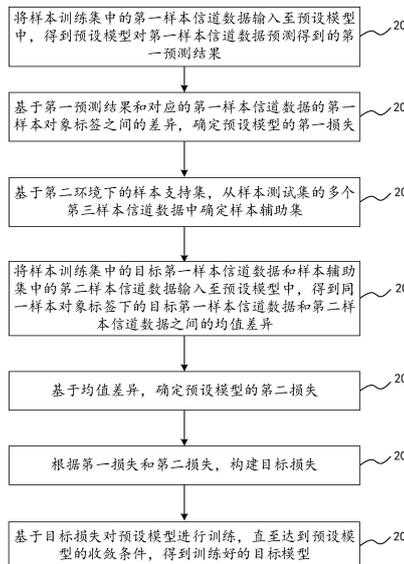
(54) 发明名称

基于少样本学习的模型对对象识别的方法及相关设备

(57) 摘要

本申请实施例提供了一种基于少样本学习的模型对对象识别的方法及相关设备。方法包括：当第二环境的信道数据的状态变化时，从第二环境中获取待分析信道数据；将待分析信道数据输入至目标模型中，得到对象标签；目标模型基于样本训练集中第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异，和基于样本训练集中的目标第一样本信道数据与在同样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定；目标对象为从多个对象中选取的对象，目标对象的数量少于对象的数量；目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。以此，能够提高模型的训练效率和识别准确性。

CN 118585901 A



1. 一种基于少样本学习的模型对对象识别的方法,其特征在于,所述方法包括:
 - 检测第二环境的信道数据的状态;
 - 当所述第二环境的信道数据的状态发生变化时,从所述第二环境中获取待分析信道数据;
 - 将所述待分析信道数据输入至目标模型中,得到对象标签;
 - 其中,所述目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于所述样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;
 - 所述样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;
 - 所述样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,所述目标对象为从所述多个对象中选取的对象,所述目标对象的数量少于所述对象的数量;
 - 所述目标第一样本信道数据和所述第二样本信道数据属于同一对象分别在所述第一环境和所述第二环境中的信道数据。
2. 根据权利要求1所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述目标模型通过以下方式训练得到:
 - 将样本训练集中的第一样本信道数据输入至预设模型中,得到所述预设模型对所述第一样本信道数据预测得到的第一预测结果;
 - 基于所述第一预测结果和对应的所述第一样本信道数据的第一样本对象标签之间的差异,确定所述预设模型的第一损失;
 - 基于第二环境下的样本支持集,从样本测试集的多个第三样本信道数据中确定样本辅助集;
 - 将所述样本训练集中的目标第一样本信道数据和所述样本辅助集中的第二样本信道数据输入至所述预设模型中,得到同一样本对象标签下的所述目标第一样本信道数据和所述第二样本信道数据之间的均值差异;
 - 基于所述均值差异,确定所述预设模型的第二损失;
 - 根据所述第一损失和所述第二损失,构建目标损失;
 - 基于所述目标损失对所述预设模型进行训练,直至达到所述预设模型的收敛条件,得到训练好的目标模型。
3. 根据权利要求2所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述基于第二环境下的样本支持集,从所述样本测试集的多个第三样本信道数据中确定样本辅助集,包括:
 - 获取第二环境下所述目标对象对应的样本支持集,以及所述样本支持集中各第四样本信道数据对应的第四样本对象标签;
 - 根据所述样本测试集中的每个所述第三样本信道数据与所述样本支持集中的多个所述第四样本信道数据的第一特征距离,确定每个所述第三样本信道数据的第三样本对象标签;
 - 从所述样本支持集中确定与所述第三样本信道数据具有相同样本对象标签的多个目

标第四样本信道数据；

根据所述多个目标第四样本信道数据计算对应的所述第三样本信道数据的评估指标；

基于每个所述第三样本信道数据对应的所述评估指标,从多个所述第三样本信道数据中确定多个所述第二样本信道数据,得到样本辅助集。

4.根据权利要求3所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述根据所述多个目标第四样本信道数据计算对应的所述第三样本信道数据的评估指标,包括:

获取每个所述第三样本信道数据与对应的多个所述目标第四样本信道数据的多个第二特征距离;

对多个所述第二特征距离进行排序,并根据得到的排序结果从所述多个第二特征距离中选取目标距离;

根据每个所述第三样本信道数据对应的目标距离,确定对应的所述第三样本信道数据的置信度;

获取每个所述第三样本信道数据的真实样本对象标签,并基于所述真实样本对象标签与所述第三样本对象标签进行比对,确定所述第三样本信道数据的第一分类准确度;

基于所述置信度和所述第一分类准确度,计算每个所述第三样本信道数据的评估指标。

5.根据权利要求2所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述将所述样本训练集中的目标第一样本信道数据和所述样本辅助集中的第二样本信道数据输入至所述预设模型中,得到同一样本对象标签下的所述目标第一样本信道数据和所述第二样本信道数据之间的均值差异,包括:

将所述样本训练集中的目标第一样本信道数据和所述样本辅助集中的第二样本信道数据输入至所述预设模型中,得到所述目标第一样本信道数据对应的第一特征向量,以及所述第二样本信道数据对应的第二特征向量;

将相同的样本对象标签下对应的所述第一特征向量和所述第二特征向量输入至所述预设模型预先设置的第二损失函数中,得到同一样本对象标签下的所述目标第一样本信道数据和所述第二样本信道数据之间的均值差异。

6.根据权利要求5所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述基于所述均值差异,确定所述预设模型的第二损失,包括:

针对每个样本对象标签,确定各样本对象标签对应的所述第一样本信道数据的第一特征向量和所述第二样本信道数据的第二特征向量的均值差异;

以样本对象标签作为划分维度,对各样本对象标签下的均值差异进行累加,得到所述预设模型的第二损失。

7.根据权利要求2所述的基于少样本学习的模型对对象识别的方法,其特征在于,所述基于所述目标损失对所述预设模型进行训练,直至达到所述预设模型的收敛条件,得到训练好的目标模型,包括:

基于所述目标损失更新所述预设模型的第一参数,并将所述样本支持集中的多个第四样本信道数据输入至所述预设模型中,得到所述多个第四样本信道数据的多个第二预测结果;

将所述第二预测结果与对应的所述第四样本对象标签进行比对,得到比对结果;
根据所述比对结果确定所述预设模型在所述第一参数下的第二分类准确度;

重复基于所述目标损失更新所述预设模型的第一参数,并将所述样本支持集中的多个第四样本信道数据输入至所述预设模型中,得到所述多个第四样本信道数据的多个第二预测结果;将所述第二预测结果与对应的所述第四样本对象标签进行比对,得到比对结果;根据所述比对结果确定所述预设模型在所述第一参数下的新的第二分类准确度,直至所述新的第二分类准确率保持不变的次数达到预设次数时,停止对所述预设模型的训练,将所述第二分类准确率最高的训练轮次对应的第一参数作为目标模型的目标参数,得到目标模型。

8. 一种基于少样本学习的模型对对象识别的装置,其特征在于,所述装置包括:

检测模块,用于检测第二环境的信道数据的状态;

获取模块,用于当所述第二环境的信道数据的状态发生变化时,从所述第二环境中获取待分析信道数据;

输入模块,用于将所述待分析信道数据输入至目标模型中,得到对象标签;其中,所述目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于所述样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;所述样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;所述样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,所述目标对象为从所述多个对象中选取的对象,所述目标对象的数量少于所述对象的数量;所述目标第一样本信道数据和所述第二样本信道数据属于同一对象分别在所述第一环境和所述第二环境中的信道数据。

9. 一种计算机设备,其特征在于,所述计算机设备包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现权利要求1至7任一项所述的基于少样本学习的模型对对象识别的方法。

10. 一种计算机可读存储介质,所述存储介质存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至7任一项所述的基于少样本学习的模型对对象识别的方法。

基于少样本学习的模型对对象识别的方法及相关设备

技术领域

[0001] 本申请涉及人工智能技术领域,尤其涉及一种基于少样本学习的模型对对象识别的方法及相关设备。

背景技术

[0002] 基于Wi-Fi的对象识别技术广泛应用于考勤、安检等对象识别领域,其能够基于Wi-Fi的对象识别技术主要利用信道状态信息(Channel State Information,CSI)进行感知,通过不同对象存在的不同行为习惯对CSI产生的不同影响,来识别当前行动的对象。以此,能够有效保护隐私,不受对象的妆发和穿着影响,同时无需排队等待。

[0003] 然而,不同的环境会导致CSI发生较大变化,使得已训练好的模型失效。而每次检测环境变化时,在新场景中采集大量带标签的CSI样本并进行重新训练是不现实的,需要耗费大量的训练资源,训练的效率低。另一方面,还存在相关技术直接将源域的训练成果应用至新的环境中,会导致模型在应用过程中对对象识别的准确率下降。

发明内容

[0004] 本申请实施例的主要目的在于提出一种基于少样本学习的模型对对象识别的方法及相关设备,能够在提高模型在少样本场景下的训练效率的同时,提高训练好的模型对对象识别的准确性。

[0005] 为实现上述目的,本申请实施例的第一方面提出了一种基于少样本学习的模型对对象识别的方法,所述方法包括:

检测第二环境的信道数据的状态;

当所述第二环境的信道数据的状态发生变化时,从所述第二环境中获取待分析信道数据;

将所述待分析信道数据输入至目标模型中,得到对象标签;

其中,所述目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于所述样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;

所述样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;

所述样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,所述目标对象为从所述多个对象中选取的对象,所述目标对象的数量少于所述对象的数量;

所述目标第一样本信道数据和所述第二样本信道数据属于同一对象分别在所述第一环境和所述第二环境中的信道数据。

[0006] 相应的,本申请实施例的第二方面提出了一种基于少样本学习的模型对对象识别的装置,所述装置包括:

检测模块,用于检测第二环境的信道数据的状态;

获取模块,用于当所述第二环境的信道数据的状态发生变化时,从所述第二环境中获取待分析信道数据;

输入模块,用于将所述待分析信道数据输入至目标模型中,得到对象标签;其中,所述目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于所述样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;所述样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;所述样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,所述目标对象为从所述多个对象中选取的对象,所述目标对象的数量少于所述对象的数量;所述目标第一样本信道数据和所述第二样本信道数据属于同一对象分别在所述第一环境和所述第二环境中的信道数据。

[0007] 在一些实施方式中,所述基于少样本学习的模型对对象识别的装置还包括训练模块,用于:

将样本训练集中的第一样本信道数据输入至预设模型中,得到所述预设模型对所述第一样本信道数据预测得到的第一预测结果;

基于所述第一预测结果和对应的所述第一样本信道数据的第一样本对象标签之间的差异,确定所述预设模型的第一损失;

基于第二环境下的样本支持集,从所述样本测试集的多个第三样本信道数据中确定样本辅助集;

将所述样本训练集中的目标第一样本信道数据和所述样本辅助集中的第二样本信道数据输入至所述预设模型中,得到同一样本对象标签下的所述目标第一样本信道数据和所述第二样本信道数据之间的均值差异;

基于所述均值差异,确定所述预设模型的第二损失;

根据所述第一损失和所述第二损失,构建目标损失;

基于所述目标损失对所述预设模型进行训练,直至达到所述预设模型的收敛条件,得到训练好的目标模型。

[0008] 在一些实施方式中,所述训练模块,还用于:

获取第二环境下所述目标对象对应的样本支持集,以及所述样本支持集中各第四样本信道数据对应的第四样本对象标签;

根据所述样本测试集中的每个所述第三样本信道数据与所述样本支持集中的多个所述第四样本信道数据的第一特征距离,确定每个所述第三样本信道数据的第三样本对象标签;

从所述样本支持集中确定与所述第三样本信道数据具有相同样本对象标签的多个目标第四样本信道数据;

根据所述多个目标第四样本信道数据计算对应的所述第三样本信道数据的评估指标;

基于每个所述第三样本信道数据对应的所述评估指标,从多个所述第三样本信道数据中确定多个所述第二样本信道数据,得到样本辅助集。

[0009] 在一些实施方式中,所述训练模块,还用于:

获取每个所述第三样本信道数据与对应的多个所述目标第四样本信道数据的多个第二特征距离;

对多个所述第二特征距离进行排序,并根据得到的排序结果从所述多个第二特征距离中选取目标距离;

根据每个所述第三样本信道数据对应的目标距离,确定对应的所述第三样本信道数据的置信度;

获取每个所述第三样本信道数据的真实样本对象标签,并基于所述真实样本对象标签与所述第三样本对象标签进行比对,确定所述第三样本信道数据的第一分类准确度;

基于所述置信度和所述第一分类准确度,计算每个所述第三样本信道数据的评估指标。

[0010] 在一些实施方式中,所述训练模块,还用于:

将所述样本训练集中的目标第一样本信道数据和所述样本辅助集中的第二样本信道数据输入至所述预设模型中,得到所述目标第一样本信道数据对应的第一特征向量,以及所述第二样本信道数据对应的第二特征向量;

将相同的样本对象标签下对应的所述第一特征向量和所述第二特征向量输入至所述预设模型预先设置的第二损失函数中,得到同一样本对象标签下的所述目标第一样本信道数据和所述第二样本信道数据之间的均值差异。

[0011] 在一些实施方式中,所述训练模块,还用于:

针对每个样本对象标签,确定各样本对象标签对应的所述第一样本信道数据的第一特征向量和所述第二样本信道数据的第二特征向量的均值差异;

以样本对象标签作为划分维度,对各样本对象标签下的均值差异进行累加,得到所述预设模型的第二损失。

[0012] 在一些实施方式中,所述训练模块,还用于:

基于所述目标损失更新所述预设模型的第一参数,并将所述样本支持集中的多个第四样本信道数据输入至所述预设模型中,得到所述多个第四样本信道数据的多个第二预测结果;

将所述第二预测结果与对应的所述第四样本对象标签进行比对,得到比对结果;

根据所述比对结果确定所述预设模型在所述第一参数下的第二分类准确度;

重复基于所述目标损失更新所述预设模型的第一参数,并将所述样本支持集中的多个第四样本信道数据输入至所述预设模型中,得到所述多个第四样本信道数据的多个第二预测结果;将所述第二预测结果与对应的所述第四样本对象标签进行比对,得到比对结果;根据所述比对结果确定所述预设模型在所述第一参数下的新的第二分类准确度,直至所述新的第二分类准确率保持不变的次数达到预设次数时,停止对所述预设模型的训练,将所述第二分类准确率最高的训练轮次对应的第一参数作为目标模型的目标参数,得到目标模型。

[0013] 相应的,本申请实施例的第三方面提出了一种计算机设备,所述计算机设备包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现本申请第一方面实施例任一项所述的基于少样本学习的模型对对象识别的方法。

[0014] 相应的,本申请实施例的第四方面提出了一种计算机可读存储介质,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时实现本申请第一方面实施例任一项所述的基于少样本学习的模型对对象识别的方法。

[0015] 本申请实施例通过检测第二环境的信道数据的状态;当第二环境的信道数据的状态发生变化时,从第二环境中获取待分析信道数据;将待分析信道数据输入至目标模型中,得到对象标签;其中,目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,目标对象为从多个对象中选取的对象,目标对象的数量少于对象的数量;目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。以此,能够按照同一样本对象标签对齐源域的样本训练集和目标域的样本辅助集,以此,能够通过中间的特征空间使得各样本对象标签在源域和目标域的分布大致相同,以使得模型在少样本训练的情况下,也能够达到良好的训练效果,提高了模型在少样本场景下的训练效率,同时由于目标模型是基于预测结果与样本对象标签之间的差异,以及目标第一样本信道数据与第二样本信道数据的均值差异确定的,由此可以提高训练好的模型对对象识别的准确性。综上,本申请能够在提高模型在少样本场景下的训练效率的同时,提高训练好的模型对对象识别的准确性。

附图说明

[0016] 图1是本申请实施例提供的基于少样本学习的模型对对象识别的系统的架构示意图;

图2是本申请实施例提供的基于少样本学习的模型对对象识别的方法的流程图;

图3是本申请实施例提供的对预设模型进行训练的步骤流程图;

图4是本申请实施例提供的预设模型的网络结构图;

图5是本申请实施例提供的对预设模型进行训练的总体流程图;

图6是本申请实施例提供的基于少样本学习的模型对对象识别的装置的功能模块示意图;

图7是本申请实施例提供的计算机设备的硬件结构示意图。

具体实施方式

[0017] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处所描述的具体实施例仅用以解释本申请,并不用于限定本申请。

[0018] 需要说明的是,虽然在装置示意图中进行了功能模块划分,在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于装置中的模块划分,或流程图中的顺序执行所示出或描述的步骤。说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0019] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同。本文中所使用的术语只是为了描述本申请实施例的目的,不是旨在限制本申请。

[0020] 基于Wi-Fi的对象识别技术广泛应用于考勤、安检等对象识别领域,其能够基于Wi-Fi的对象识别技术主要利用信道状态信息(Channel State Information,CSI)进行感知,通过不同对象存在的不同行为习惯对CSI产生的不同影响,来识别当前行动的对象。以此,能够有效保护隐私,不受对象的妆发和穿着影响,同时无需排队等待。

[0021] 然而,不同的环境会导致CSI发生较大变化,使得已训练好的模型失效。而每次检测环境变化时,在新场景中采集大量带标签的CSI样本并进行重新训练是不现实的,需要耗费大量的训练资源,训练的效率低。

[0022] 另一方面,还存在相关技术直接将源域的训练成果应用至新的环境中,会导致模型在应用过程中对对象识别的准确率下降。具体来说,传统域适应方法基于假设 $P_t(x|y)=P_s(x|y)$,但实际上, $P_t(x) \neq P_s(x)$,这会导致神经网络难以学习到使得 $P_t(y|x;\theta)=P_s(y|x;\theta)$ 的正确参数,因此, $P_t(x|y)=P_s(x|y)$ 的假设在大部分情况下是不成立的,也就是说,相关技术中使整个源域和目标域的分布相同是不合理的。具体可以通过条件概率公式构建得到:

$$\begin{aligned}
 P_s(y|x) &= P_t(y|x) \frac{P_s(x|y)P_s(y)}{P_s(x)} \frac{P_t(x)}{P_t(x|y)P_t(y)} \\
 [0023] \quad &= P_t(y|x) \frac{P_t(x)}{P_s(x)} \frac{P_s(x|y)}{P_t(x|y)} \frac{P_s(y)}{P_t(y)} \\
 [0024] \quad &= P_t(y|x) \frac{\sum_{y'} P_t(y')P_t(x|y')}{\sum_{y'} P_s(y')P_s(x|y')} \frac{P_s(x|y)}{P_t(x|y)} \frac{P_s(y)}{P_t(y)}
 \end{aligned}$$

[0025] 在实际的场景中,有 $P_t(y) \approx P_s(y)$,即各类别样本是平衡的,在上述公式的基础上,如果 $P_t(x|y) \approx P_s(x|y)$ 成立,就可以使得 $P_t(y|x) \approx P_s(y|x)$ 。因此,本申请通过寻找中间的特征空间,使得在各类别中源域和目标域的分布大致相同,而不是像相关技术中希望使得整个源域和目标域的分布相同。

[0026] 基于此,本申请实施例提供了一种基于少样本学习的模型对对象识别的方法及相关设备,能够在提高模型在少样本场景下的训练效率的同时,提高训练好的模型对对象识别的准确性。

[0027] 本申请实施例提供的基于少样本学习的模型对对象识别的方法及相关设备,具体通过如下实施例进行说明,首先描述本申请实施例中的基于少样本学习的模型对对象识别的系统。

[0028] 请参照图1,在一些实施方式中,本申请实施例提供一种基于少样本学习的模型对对象识别的系统,基于少样本学习的模型对对象识别的系统包括终端11和服务端12。

[0029] 具体的,终端11可以是用于数据采集和初步处理的设备。例如,终端11可以是带单天线的ESP32-S3接收机,用于采集CSI数据。终端11可以实时检测或者定时检测目标环境的

CSI数据的变化,当检测到CSI数据发生变化时,对采集到的数据进行预处理,并将采集到的CSI数据传输到服务器端进行进一步处理。

[0030] 在一些实施方式中,服务器端12是进行数据处理、模型训练和对象识别的计算平台,可以进行数据的高效处理、模型训练和参数优化,以及最终模型性能的测试和评估。服务器端12接收到来自终端11的CSI数据后,可以对CSI数据进行进一步的处理,如划分样本和构建样本训练集、样本测试集、样本支持集和样本辅助集,并通过样本训练集和样本辅助集对预设模型进行训练,使得训练得到的目标模型具备对新的环境的对象进行准确识别的能力;或者,在得到目标模型之后,可以对CSI数据进行对象的直接识别。

[0031] 本申请实施例中的基于少样本学习的模型对对象识别的方法可以通过如下实施例进行说明。

[0032] 需要说明的是,在本申请的各个具体实施方式中,当涉及到需要根据用户信息、用户行为数据,用户历史数据以及用户位置信息等与用户身份或特性相关的数据进行相关处理时,都会先获得用户的许可或者同意。而且,对这些数据的收集、使用和处理等,都会遵守相关法律法规和标准。此外,当本申请实施例需要获取用户的敏感个人信息时,会通过弹窗或者跳转到确认页面等方式获得用户的单独许可或者单独同意,在明确获得用户的单独许可或者单独同意之后,再获取用于使本申请实施例能够正常运行的必要的用户相关数据。

[0033] 在本申请实施例中,将从基于少样本学习的模型对对象识别的装置的维度进行描述,该基于少样本学习的模型对对象识别的装置具体可以集成在计算机设备中。参见图2,图2为本申请实施例提供的基于少样本学习的模型对对象识别的方法的步骤流程图,本申请实施例以基于少样本学习的模型对对象识别的装置具体集成在如终端或服务器上为例,终端或服务器上的处理器执行基于少样本学习的模型对对象识别的方法对应的程序指令时,具体流程如下:

步骤101,检测第二环境的信道数据的状态。

[0034] 可以理解的是,为了能够准确在出现识别对象时,及时、准确地对对象进行识别,可以检测对象所在的第二环境的信道数据的状态,以确定是否需要采集对应的信道数据。

[0035] 其中,第二环境可以是进行对象识别的特定环境或者场景。例如,在对人员打卡进行识别时,人员打卡时所处的区域为第二环境;进行安检时,安检区域为第二环境。第二环境可以是多个不同于第一环境的新的环境,对不同的第二环境,均可以采用本申请的方式进行处理。进一步的,第二环境可以根据实际情况进行调整,只需要确保通过第二环境能够确保采集到待识别对象的待分析信道数据即可。

[0036] 其中,信道数据可以是在无线通信系统中,由无线信道的特性所产生的一系列参数或特征。信道数据可以包括信号的幅度、相位、时延等,可以用于表征无线信号在传输过程中的变化。

[0037] 在一些实施方式中,可以对第二环境的信道数据的状态进行定时检测或者实时检测。定时检测设置的时间间隔可以为5秒、10秒等等。进一步的,可以根据不同的时间段自适应调整检测的方式,例如在高峰期可以采用实时检测,在少人阶段可以采用定时检测,定时检测的时间间隔可自行设置。

[0038] 示例性的,信道数据的状态可以通过无线传感器、基站或移动设备等收集。

[0039] 通过检测第二环境的信道数据的状态,可以便于后续确定信道数据的状态是否发

生变化,以便于在信道数据的状态发生变化时,获取待分析信道数据。

[0040] 步骤102,当第二环境的信道数据的状态发生变化时,从第二环境中获取待分析信道数据。

[0041] 在一些实施方式中,当第二环境的信道数据的状态发生变化时,表明第二环境中可能出现新的对象进入,此时需要对新的对象进行识别,以准确、及时地识别新的对象。

[0042] 其中,待分析信道数据可以是待识别的对象对应的信道数据,可以包括信道的幅度、相位、时延等,用于反映无线信号在传播过程中的特性。

[0043] 具体的,当环境中新的对象(如人员)进入时,会对无线信号造成干扰和遮挡,改变信号的强度和传播路径,从而改变信道数据的状态。因此,可以通过信道数据的状态是否发生改变,来判断第二环境中是否有新的人员进入,以更加准确地检测和识别对象。

[0044] 通过从第二环境中获取待分析信道数据,可以便于后续对进入的对象进行及时、准确的识别。

[0045] 步骤103,将待分析信道数据输入至目标模型中,得到对象标签。

[0046] 其中,目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于样本训练集中的目标第一样本信道数据与在同一对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;

样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;

样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,目标对象为从多个对象中选取的对象,目标对象的数量少于对象的数量;

目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。

[0047] 在一些实施方式中,为了对对象进行识别,提高场所的安全性以及人员识别的效率,可以将对象对应的待分析信道数据输入至已经经过预先训练的目标模型中,得到待分析对象的对象标签,也即具体的类别。以此,可以快速、准确地对对象进行识别。

[0048] 其中,目标模型可以是用于识别对象标签的机器学习模型,例如支持向量机、卷积神经网络、循环神经网络等等。

[0049] 其中,对象标签可以是待分析对象的具体类别,通过将待分析信道数据输入目标模型,可以直接得到对应的对象标签。示例性的,对象标签可以是人员A。

[0050] 其中,样本训练集可以是用于训练目标模型的数据集,通过将样本训练集中的第一样本信道数据输入至模型进行训练,并结合对应的第一样本对象标签对模型的参数进行调整,可以使得模型学习数据的特征以及这些特征与标签之间的关系。

[0051] 其中,第一样本信道数据可以是样本训练集中的样本信道数据。例如,第一样本信道数据可以是用户A经过办公楼(第一环境)时的信道数据、用户B经过第一环境时的信道数据,等等。

[0052] 其中,第一预测结果可以是将第一样本信道数据输入至预设模型之后,预设模型输出的预测对象标签,第一预测结果可以用于反馈当前模型的性能。例如,对第一样本信道数据对应的第一预测结果为用户A。

[0053] 其中,第一样本对象标签可以是第一样本信道数据的真实对象标签,用于与第一预测结果进行差异比对,以检验预设模型的分类性能,例如,第一样本信道数据对应的第一样本对象标签为用户B。

[0054] 其中,目标第一样本信道数据可以是在第一环境中目标对象的信道数据,目标第一样本信道数据所属的目标对象在第二环境中,有与之对应的第二样本信道数据。

[0055] 其中,样本辅助集可以是从小样本测试集中选取的评估指标较高的样本信道数据所构成的数据集,用于使得预设模型在第一环境和第二环境中进行样本对象标签的对齐。

[0056] 其中,第二样本信道数据可以是在第二环境中从目标对象获取的信道数据。第二样本信道数据所属的目标对象在第一环境中,有与之对应的目标第一样本信道数据。

[0057] 其中,第一环境可以是采集第一样本信道数据时的环境。可以理解的是,由于信道状态信息是对无线信号在不同环境中传播特性的描述,而这些特性会随着环境的变化而变化。例如,物体的位置、墙壁的材料、人员的移动等因素都会影响无线信号的传播路径和强度。因此,在跨环境应用中,需要考虑环境变化对目标对象的信道数据的影响,以确保目标模型在新环境中也具有较佳表现,因此,除了采用第一环境的第一样本信道数据对预设模型进行训练,还需要将第二环境中的第二样本信道数据与第一环境的目标样本信道数据进行样本对象标签的对齐。

[0058] 其中,第二样本对象标签可以是第二样本信道数据对应的真实标签。

[0059] 示例性的,若要训练得到目标模型来识别员工在不同环境下的身份,可以通过以下方式对预设模型进行训练,在不违背本申请构思的前提下,以下示例可应用至其他环境或者场景。首先,若第一环境为办公楼,已经收集用户A、B、C等员工经过办公楼时的第一样本信道数据。第一样本信道数据构成了样本训练集,并有对应的第一样本对象标签,如用户A的第一样本对象标签为A,用户B的为B,依此类推。

[0060] 进一步的,若第二环境为公司门禁,此时确定公司门禁需要识别的目标对象,例如为用户A和用户B,在公司门禁的第二样本信道数据,第二样本信道数据构成了样本辅助集,并具有对应的第二样本对象标签。

[0061] 此时,将用户A在办公楼中的信道数据输入预设模型,预设模型输出的第一预测结果可能为A或其他标签。根据第一预测结果与第一样本对象标签来确定第一损失。进一步的,可以对比用户A在办公楼与公司门禁的样本信道数据,利用两者的特征值的均值差异构建第二损失,以通过第二损失使得预设模型能够基于样本对象标签A进行用户A对应的目标第一样本信道数据和第二样本信道数据的对齐,以确保在公司门禁的样本量较少的情况下,也能够利用从样本训练集中得到的经验,使得预设模型能够在公司门禁中也能准确识别用户A。

[0062] 进一步的,可以通过第一损失和第二损失构建目标损失对预设模型进行训练,并在训练的过程中不断调整模型参数,使得模型能够从第一环境的已有的样本训练集中学习样本信道数据的特征与对应的样本对象标签的关系,并实现在第二环境中的应用,使得目标模型在第二环境中,也能够实现对第二样本信道数据的准确识别。

[0063] 本申请实施例通过检测第二环境的信道数据的状态;当第二环境的信道数据的状态发生变化时,从第二环境中获取待分析信道数据;将待分析信道数据输入至目标模型中,得到对象标签;其中,目标模型基于样本训练集中每个第一样本信道数据的第一预测结果

与对应的第一样本对象标签之间的差异,以及基于样本训练集中的目标第一样本信道数据与在同样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,目标对象为从多个对象中选取的对象,目标对象的数量少于对象的数量;目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。以此,能够按照同样本对象标签对齐源域的样本训练集和目标域的样本辅助集,以此,能够通过中间的特征空间使得各样本对象标签在源域和目标域的分布大致相同,以使得模型在少样本训练的情况下,也能够达到良好的训练效果,提高了模型在少样本场景下的训练效率,同时由于目标模型是基于预测结果与样本对象标签之间的差异,以及目标第一样本信道数据与第二样本信道数据的均值差异确定的,由此可以提高训练好的模型对对象识别的准确性。综上,本申请能够在提高模型在少样本场景下的训练效率的同时,提高训练好的模型对对象识别的准确性。

[0064] 请参照图3,在一些实施方式中,为了使得模型能够在新的环境只存在少样本信道数据的情况下,依然能够快速适应新的环境并对新环境的样本信道数据作出准确识别,可以通过对预设模型进行训练,使得训练好的目标模型既具备准确的分类能力,也能够具有对新环境下的良好适应性。例如,目标模型通过以下方式训练得到:

步骤201,将样本训练集中的第一样本信道数据输入至预设模型中,得到预设模型对第一样本信道数据预测得到的第一预测结果。

[0065] 在一些实施方式中,可以将第一样本信道数据输入至预设模型中,预设模型可以根据当前的模型参数对输入的数据进行处理,并输出第一预测结果,以便于后续根据第一预测结果与第一样本信道数据的第一样本对象标签进行比较,确定在当前的模型参数下,预设模型的分性能,以便于后续对预设模型进行改进。示例性的,第一样本信道数据为用户A对应的样本信道数据,得到的第一预测结果可能为用户A,也可能为用户B,也即是说,预设模型对于第一样本信道数据的第一预测结果可能并不准确。

[0066] 步骤202,基于第一预测结果和对应的第一样本信道数据的第一样本对象标签之间的差异,确定预设模型的第一损失。

[0067] 其中,第一损失可以是交叉熵损失,用于衡量预设模型对于第一样本信道数据的第一预测结果(预测值)与第一样本对象标签(真实值)之间的差异的函数,也即是第一预测结果与第一样本对象标签之间差异的量化指标。

[0068] 具体的,第一损失越低,表明预设模型对第一样本信道数据的预测越准确,通过反向传播或者其他优化算法,预设模型可以根据第一损失的值调整模型参数,以减少损失。

[0069] 通过以上方式,可以确定预设模型的第一损失,以便于后续基于第一损失与第二损失结合,提高预设模型的分性能。

[0070] 步骤203,基于第二环境下的样本支持集,从样本测试集的多个第三样本信道数据中确定样本辅助集。

[0071] 其中,样本支持集可以是在第二环境中的样本信道数据,样本支持集可以用于从样本测试集中筛选样本辅助集,样本支持集还可以用于在预设模型的训练过程中进行早停决策,即根据预设模型在支持集上的准确率变化来决定何时对预设模型停止训练。

[0072] 其中,样本测试集可以是第二环境中的数据,可以用于选取质量较高的第二样本信道数据。

[0073] 其中,第三样本信道数据可以是样本测试集上的样本信道数据,第三样本信道数据与第二样本信道数据、待分析信道数据都处于第二环境中。

[0074] 在一些实施方式中,为了避免样本支持集的质量影响预设模型在第二环境中的分类准确度,可以基于样本支持集对样本测试集中的第三样本信道数据进行分类,并确定各第三样本信道数据的评估指标,以从多个第三样本信道数据中选取较高评估指标的第二样本信道数据对预设模型进行训练。

[0075] 通过以上方式,可以从样本测试集的多个第三样本信道数据中确定质量较高、分类特征更加明显的样本辅助集,以便于后续确定同一样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异。

[0076] 在一些实施方式中,为了提高目标模型在跨环境应用中的准确性和稳定性,克服传统方法的局限性,实现更高效的跨域识别,可以通过构建样本辅助集,引入高置信度样本,使得预设模型在特征空间的对齐过程更加稳定和准确。例如,步骤203可以包括:

(203.1) 获取第二环境下目标对象对应的样本支持集,以及样本支持集中各第四样本信道数据对应的第四样本对象标签;

(203.2) 根据样本测试集中的每个第三样本信道数据与样本支持集中的多个第四样本信道数据的第一特征距离,确定每个第三样本信道数据的第三样本对象标签;

(203.3) 从样本支持集中确定与第三样本信道数据具有相同样本对象标签的多个目标第四样本信道数据;

(203.4) 根据多个目标第四样本信道数据计算对应的第三样本信道数据的评估指标;

(203.5) 基于每个第三样本信道数据对应的评估指标,从多个第三样本信道数据中确定多个第二样本信道数据,得到样本辅助集。

[0077] 其中,第四样本信道数据可以是样本支持集中的样本信道数据。

[0078] 其中,第四样本对象标签可以是第一样本信道数据的真实样本对象标签。

[0079] 其中,第一特征距离可以是第三样本信道数据与第四样本信道数据的欧氏距离。示例性的,第一特征距离还可以是曼哈顿距离、马氏距离、余弦相似度等等。

[0080] 其中,第三样本对象标签可以通过第一特征距离确定的第三样本信道数据的样本对象标签。由于原样本测试集中的第三样本对象是不具有样本对象标签的,因此,可以基于样本支持集中各第四样本信道数据的第四样本对象标签确定第三样本信道数据的第三样本对象标签。

[0081] 其中,目标第四样本信道数据可以是与对应的第三样本信道数据具有相同的样本对象标签的第四样本信道数据。例如第三样本信道数据的第三样本对象标签为A,若第四样本信道数据的第四样本对象标签也为A,那么该第四样本信道数据为目标样本信道数据。

[0082] 其中,评估指标可以是第三样本信道数据的置信度和准确度的综合指标,用于评估第三样本信道数据的质量。进一步的,评估指标也可以是置信度和准确度中的一个。

[0083] 示例性的,当第一特征距离为欧氏距离时,可以通过计算每个第三样本信道数据与样本支持集中的所有第四样本信道数据的欧氏距离,来确定每个第三样本信道数据的第

三样本对象标签。示例性的,针对每个第三样本信道数据,可以选取欧氏距离最近的k个第四样本信道数据,并获取k个第四样本信道数据的第四样本对象标签,将k个第四样本信道数据中出现频率最高第四样本对象标签作为第三样本信道数据的第三样本对象标签。进一步的,k的值可以根据实际情况确定。

[0084] 例如,当与第三样本信道数据的欧式距离最近的k个第四样本信道数据为信道数据A,对应的第四样本对象标签为1、信道数据B,对应的第四样本对象标签为1、信道数据C,对应的第四样本对象标签为1、信道数据D,对应的第四样本对象标签为2、信道数据E,对应的第四样本对象标签为1,则第三样本信道数据的第三样本对象标签为1。

[0085] 进一步的,可以基于第三样本信道数据对应的第三样本对象标签,从样本支持集中确定相同样本对象标签的第四样本信道数据。例如,第三样本信道数据的第三样本对象标签为A,那么可以从样本支持集中确定样本对象标签为A的所有第四样本信道数据作为目标第四样本信道数据。

[0086] 进一步的,可以根据每个第三样本信道数据对应的多个目标第四样本信道数据确定第三样本信道数据的置信度和对第三样本信道数据分类的准确度,并根据置信度和准确度来计算第三样本信道数据的评估指标。进一步的,也可以只将置信度或者准确度作为第三样本信道数据的评估指标。

[0087] 示例性的,可以从第三样本信道数据中选取评估指标较高的第二样本信道数据,并根据多个第二样本数据构建样本辅助集。

[0088] 通过特征距离匹配和评估指标的计算,能够有效确定各第三样本信道数据的第三样本对象标签,并从中选取质量较高的第二样本信道数据,构建样本辅助集,以通过利用样本支持集中的标记数据,提供了一种系统化的途径来分类和优化测试数据,使得结果更加可靠和精确。

[0089] 在一些实施方式中,为了构建可靠的样本辅助集,减少错误样本的影响,可以通过计算各第三样本信道数据的评估指标,有效地选择和利用高质量样本,从而提高预设模型在不同域环境下的表现。例如,(203.4)包括:

(203.4.1) 获取每个第三样本信道数据与对应的多个目标第四样本信道数据的多个第二特征距离;

(203.4.2) 对多个第二特征距离进行排序,并根据得到的排序结果从多个第二特征距离中选取目标距离;

(203.4.3) 根据每个第三样本信道数据对应的目标距离,确定对应的第三样本信道数据的置信度;

(203.4.4) 获取每个第三样本信道数据的真实样本对象标签,并基于真实样本对象标签与第三样本对象标签进行比对,确定第三样本信道数据的第一分类准确度;

(203.4.5) 基于置信度和第一分类准确度,计算每个第三样本信道数据的评估指标。

[0090] 其中,第二特征距离可以是第三样本信道数据与目标第四样本信道数据在特征空间中的距离,用于评估第三样本信道数据与第四样本信道数据之间的相似性。例如,第二特征距离可以是欧氏距离、曼哈顿距离、马氏距离或余弦相似度等等。

[0091] 其中,排序结果可以是根据第三样本信道数据与第四样本信道数据之间的第二特

征距离进行排序得到的结果,排序方式可以为顺序排序或者逆序排序。

[0092] 其中,目标距离可以从排序结果中,选取最小的第二特征距离。例如,第二特征距离按照逆序排序可以得到4、3、3、2、1,那么1为目标距离。

[0093] 其中,置信度可以是一个衡量第三样本信道数据标签可靠性的指标,置信度可以基于目标距离进行计算,具体为目标距离的倒数,例如目标距离为2,则置信度为1/2。

[0094] 其中,真实样本对象标签可以是第三样本信道数据的真实标签,第三样本信道数据的真实样本对象标签可以人为确定,也可以基于可靠的技术手段确定。

[0095] 其中,第一分类准确度可以是对第三样本信道数据的分类的准确度,可以基于第三样本对象标签与真实样本对象标签之间的匹配度进行计算。

[0096] 示例性的,获取每个第三样本信道数据与对应的多个目标第四样本信道数据的多个第二特征距离

例如,若有一个第三样本信道数据T3和三个目标第四样本信道数据T41、T42、T43,通过计算T3分别和T41、T42、T43之间的欧氏距离(可以替换为其他距离度量,如曼哈顿距离等),得到以下结果: $d(T3, T41)=5$ 、 $d(T3, T42)=3$ 、 $d(T3, T43)=8$,目标距离为最小的第二特征距离,即3。进一步的,置信度确定为1/3。

[0097] 进一步的,可以获取T3的真实样本对象标签,若T3的真实样本对象标签为A,第三样本对象标签也是A,则第三样本信道数据的第一分类准确率为1(或100%)。进一步的,可以结合置信度和第一分类准确度计算评估指标,具体的评估指标可以是置信度与第一分类准确度的乘积,此处示例为1/3。

[0098] 示例性的,可以基于每个第三样本信道数据对应的评估指标,确定预设个数或者预设比例的第三样本信道数据,具体选取规则是选取较高评估指标的第三样本信道数据,预设个数和预设比例可以根据实际情况进行设定。例如,若预设个数为15,那么,可以选取评估指标最高的15个第三样本信道数据作为第二样本信道数据,并基于这15个第三样本信道数据构建样本辅助集。

[0099] 通过结合置信度与分类准确度,评估第三样本信道数据的可靠性,选取评估指标较高的样本构建样本辅助集,能够有效提高样本信道数据选择的精准性和可靠性,从而增强模型的整体性能。

[0100] 步骤204,将样本训练集中的目标第一样本信道数据和样本辅助集中的第二样本信道数据输入至预设模型中,得到同一样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异。

[0101] 在一些实施方式中,为了使得各样本对象标签(即各类别)中源域和目标域的分布大致相同,可以确定源域(样本训练集)和目标域(样本辅助集)的中间特征空间,以更精确地对齐源域和目标域中的类别分布,进一步提高目标模型的域适应效果。

[0102] 其中,预设模型可以是待训练的模型,对预设模型进行训练,预设模型收敛之后,可以得到目标模型。

[0103] 其中,均值差异可以是目标第一样本信道数据和第二样本信道数据在相同样本对象标签下的均值差异。

[0104] 请参照图4,图4为预设模型的网络结构。示例性的,针对每个样本对象标签,可以将同一样本对象标签的目标第一样本信道数据和第二样本信道数据输入至Resnet(残差)

网络中,Resnet网络会输出目标第一样本信道数据和第二样本信道数据的特征向量,之后将同一样本对象标签的特征向量输入至去边际化核最大均值差异(Marginalized Kernel Maximum Mean Discrepancy,MK-MMD)损失函数中,计算目标第一样本信道数据和第二样本信道数据在相同样本对象标签下的均值差异。

[0105] 通过以上方式,可以得到同一样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异,以便于后续确定预设模型的第二损失。

[0106] 在一些实施方式中,为了有效减少同一样本对象标签下不同信道数据之间的分布差异,提高目标模型在第二环境的泛化能力,可以确定同一样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异,以使得同一样本对象标签下的目标第一样本信道数据和第二样本信道数据实现在中间特征空间的对齐。例如,步骤204可以包括:

(204.1) 将样本训练集中的目标第一样本信道数据和样本辅助集中的第二样本信道数据输入至预设模型中,得到目标第一样本信道数据对应的第一特征向量,以及第二样本信道数据对应的第二特征向量;

(204.2) 将相同的样本对象标签下对应的第一特征向量和第二特征向量输入至预设模型预先设置的第二损失函数中,得到同一样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异。

[0107] 其中,第一特征向量可以是在预设模型的特征提取阶段产生的向量,代表了目标第一样本信道数据在中间特征空间中的表示。第一特征向量可以用于表征目标第一样本信道数据的本质属性。

[0108] 其中,第二特征向量可以是在预设模型的特征提取阶段产生的向量,代表了第二样本信道数据在中间特征空间中的表示。第一特征向量可以用于表征第二样本信道数据的本质属性。

[0109] 其中,第二损失函数可以是MK-MMD损失函数。第二损失函数被用于每个样本对象标签下的目标第一样本信道数据和第二样本信道数据之间的均值差异。

[0110] 示例性的,可以将具有相同样本对象标签的目标第一样本信道数据和第二样本信道数据输入至预设模型中,例如,对于样本对象标签A,可以从样本训练集中提取多个目标第一样本信道数据,从样本辅助集中提取多个第二样本信道数据,并将多个目标第一样本信道数据和多个第二样本信道数据同时输入至预设模型中,预设模型的Resnet网络将会输出目标第一样本信道数据的第一特征向量和第二样本信道数据的第二特征向量。例如,多个目标第一样本信道数据对应的第一特征向量为 $[0.5, 0.3, 0.2]$,多个第二样本信道数据对应的第二特征向量为 $[0.4, 0.35, 0.25]$ 。

[0111] 进一步的,通过MK-MMD损失函数,可以计算第一特征向量和第二特征向量之间的均值差异。示例性的,可以计算第一特征向量和第二特征向量在中间特征空间的均值嵌入之间的差异,或者最大均值差异。例如,可以是第一特征向量和得让特征向量在最大均值差异(Maximum Mean Discrepancy,RKHS)中均值嵌入的差异。

[0112] 通过计算每个样本对象标签对应的第一特征向量和第二特征向量之间的均值差异,可以使得预设模型能够识别并减少不同数据源(例如样本训练集和样本辅助集)之间的不一致性,提高预设模型在未见过的第二环境中的泛化能力。

[0113] 步骤205,基于均值差异,确定预设模型的第二损失。

[0114] 在一些实施方式中,为了帮助预设模型学会如何使相同样本对象标签的不同样本信道数据在特征空间中尽可能相似,可以基于均值差异确定预设模型的第二损失,以使得样本训练集和样本辅助集中属于同一样本对象标签的样本信道数据实现在类别上的对齐,进一步提高目标模型的泛化能力以及分类的准确性。

[0115] 其中,第二损失可以基于多个样本对象标签对应的均值差异累加得到。

[0116] 示例性的,若有样本对象标签1和样本对象标签2。针对样本对象标签1,可以计算得到多个目标第一样本信道数据的第一特征向量和多个第二样本信道数据的第二特征向量对应的均值差异1;针对样本对象标签2,可以计算得到多个目标第一样本信道数据的第一特征向量和多个第二样本信道数据的第二特征向量对应的均值差异2;将均值差异1和均值差异2相加,可以得到预设模型的第二损失。

[0117] 通过引入第二损失,样本训练集和样本辅助集中属于相同样本对象标签的不同样本信道数据在特征空间中的相似性得以增强,这种类别上的对齐有助于预设模型更好地学习不同样本信道数据的共性,进而提升目标模型在处理未知数据时的表现,使得预设模型在少样本情况下也能实现高效学习。

[0118] 在一些实施方式中,为了提高目标模型的泛化能力,可以针对样本辅助集中的每个第二样本对象标签的第二样本信道数据与具有相同样本对象标签的样本训练集中的第一样本信道数据实现中间特征空间的对齐,为了步骤205可以包括:

(205.1) 针对每个样本对象标签,确定各样本对象标签对应的第一样本信道数据的第一特征向量和第二样本信道数据的第二特征向量的均值差异;

(205.2) 以样本对象标签作为划分维度,对各样本对象标签下的均值差异进行累加,得到预设模型的第二损失。

[0119] 在一些实施方式中,可以针对每个样本对象标签,获取样本辅助集的多个第二样本信道数据,以及相同样本对象标签下,样本训练集的多个目标第一样本信道数据,并将多个第二样本信道数据和多个目标第一样本信道数据同时输入至预设模型的Resnet网络,得到多个目标第一样本信道数据对应的第一特征向量和多个第二样本信道数据对应的第二特征向量,之后,计算该样本对象标签下,第一特征向量和第二特征向量的均值差异。

[0120] 进一步的,可以样本对象标签作为划分维度,也即不同样本对象标签对应不同的维度,每个维度对应一个均值差异,也即一个样本对象标签对应一个均值差异,将所有维度的均值差异相加,即可得到预设模型的第二损失。

[0121] 通过计算预设模型的第二损失,以上方式,计算MK-MMD损失函数的有益效果在于,当各类别样本是平衡的且源域和目标域的条件分布相似时,它可以在特征空间中确保不同类别的源域和目标域样本的分布一致。这比传统方法希望在特征空间中使整体源域和目标域的分布相同的方法更有效。通过这种方式,可以提高跨域任务中模型的性能,特别是在源域和目标域条件分布相近的情况下。

[0122] 步骤206,根据第一损失和第二损失,构建目标损失。

[0123] 在一些实施方式中,为了更好地解决目标模型在新的环境适应中的分布对齐问题,可以通过将第一损失(交叉熵损失)和第二损失(MK-MMD损失)相结合来进行网络训练,使得目标模型在最小化分布差异的同时,确保分类性能。

[0124] 请参照图4,在一些实施方式中,当第一损失函数为交叉熵损失函数,第二损失函

数为MK-MMD损失函数时,针对同一样本对象标签下的特征向量,可以使用样本训练集中的第一样本信道数据训练预设模型,并通过交叉熵损失函数计算第一损失。

[0125] 同时,针对Resnet网络提取出的特征空间添加MK-MMD损失函数。具体来说,在计算第二损失时,需要针对不同的样本对象标签分别计算第二损失,即对于样本训练集和样本辅助集中不同样本对象标签的样本信道数据构建不同的MK-MMD损失(即第二损失)。最终的目标损失公式为:

$$L = L_{cls} + \sum_i MKMMD(x_i^s, x_i^t)$$

[0126] 其中, L_{cls} 为交叉熵损失,即第一损失, $\sum_i MKMMD(x_i^s, x_i^t)$ 为第二损失, x_i^s 代表样本训练集中第二样本对象标签为i的全部第二样本信道数据集, x_i^t 代表样本辅助集中第一样本对象标签为i的全部目标第一样本信道数据集。

[0127] 通过第一损失可以使得目标模型具备更好的分类性能,通过第二损失可以最小化源域和目标域的分布差异,将第一损失和第二损失相加,得到目标损失,可以便于后续根据目标损失对预设模型进行训练,从而提高训练的效率。

[0128] 步骤207,基于目标损失对预设模型进行训练,直至达到预设模型的收敛条件,得到训练好的目标模型。

[0129] 在一些实施方式中,可以基于第一损失和第二损失联合得到的目标损失对预设模型进行训练,并在此过程中优化预设模型的第一参数,直至达到收敛条件,即可停止对预设模型的训练,得到训练好的目标模型。

[0130] 示例性的,收敛条件可以是对预设模型进行训练的次数达到了预设训练次数,例如达到了预设的100次;在一些实施方式中,可以在每次对预设模型的第一参数进行调整之后,将样本支持集中的第四样本信道数据输入至调整了第一参数的预设模型中,将如图4的线性层输出的第二预测结果与第四样本信道数据的第四样本对象标签进行比较,并计算分类的准确度,当分类的准确度在连续次数不变时,例如分类的准确度达到30次不发生变化时,即可停止对预设模型的训练,得到目标模型。

[0131] 通过以上方式,可以使得训练好的目标模型在第二环境中,具备良好的分类能力,同时提高了目标模型的性能。

[0132] 在一些实施方式中,为了解决相关技术中的少样本学习方法无法通过早停寻找最优参数等问题,可以利用样本支持集数据设计早停方案,通过预设模型在样本支持集上的准确度确定最佳模型参数,以有效避免过拟合,确保了目标模型的泛化能力。示例性的,步骤207可以包括:

基于目标损失更新预设模型的第一参数,并将样本支持集中的多个第四样本信道数据输入至预设模型中,得到多个第四样本信道数据的多个第二预测结果;

将第二预测结果与对应的第四样本对象标签进行比对,得到比对结果;

根据比对结果确定预设模型在第一参数下的第二分类准确度;

重复基于目标损失更新预设模型的第一参数,并将样本支持集中的多个第四样本信道数据输入至预设模型中,得到多个第四样本信道数据的多个第二预测结果;将第二预测结果与对应的第四样本对象标签进行比对,得到比对结果;根据比对结果确定预设模型

在第一参数下的新的第二分类准确度,直至新的第二分类准确率保持不变的次数达到预设次数时,停止对预设模型的训练,将第二分类准确率最高的训练轮次对应的第一参数作为目标模型的目标参数,得到目标模型。

[0133] 其中,第二预测结果可以是预设模型在接收到多个第四样本信道数据后,进行预测后的预测输出,第二预测结果是对第四样本信道数据所属类别的判断。

[0134] 其中,比对结果可以是将预设模型的第二预测结果与对应的第四样本对象标签进行比较得到的结果,比对结果可以为相同或者不相同,以此可以评估预设模型的预测准确性。

[0135] 其中,第二分类准确度可以是预设模型在预设参数设置下正确分类的准确率,用于表征预设模型预测正确的样本信道数据与总样本信道数据的数量的比例。

[0136] 其中,第一参数可以是预设模型内部的变量,决定了预设模型的具体形式和决策边界,第一参数可以通过目标损失进行更新。

[0137] 其中,预设次数可以是预设模型设定的第二分类准确度不发生变化的次数,预设次数可以根据实际情况进行调整,例如预设次数可以为20次、30次等等。

[0138] 示例性的,当有10个第四样本信道数据,比对结果表征预设模型进行准确分类的第四样本信道数据有6个,则第二分类准确度正确率为60%。

[0139] 示例性的,在预设模型每更新一次第一参数之后,均将样本支持集的多个第四样本信道数据输入至更新了第一参数的预设模型中,并计算预设模型对多个第四样本信道数据的第二分类准确度。并将第二分类准确度与上一次调整了第一参数后计算的第三分类准确度进行比较,记录第二分类准确度连续不变的次数,当第二分类准确度连续不变的次数超过预设次数时,即可停止对预设模型的训练,并将历史的第二分类准确度对应的第一参数作为目标模型的目标参数,得到训练好的目标模型。

[0140] 请参照图5,在一些实施方式中,结合图5对本申请的总体训练过程的实施例进行介绍。

[0141] 示例性的,可以使用带单天线的ESP32-S3作为接收机,采样率设为100Hz,连接家用Wi-Fi路由器采集样本信道数据。首先,让8名志愿者在环境1中采集6分钟样本信道数据并标注构建样本训练集。然后,让8名志愿者在环境2中采集3秒数据并标注构建样本支持集。接下来,让8名志愿者随意进出环境2(每次仅1人进入环境),采集约6分钟数据构建样本测试集(无需记录具体进入人员,也即无需标注对应的样本对象标签)。

[0142] 进一步的,可以将所有样本信道数据的幅度值作为预设模型的输入,并使用插值方法来填补丢失的数据。然后,使用不重叠的滑动窗口将输入数据划分为多个1秒长的样本。因此,样本支持集中的3秒数据对应于3个第四样本信道数据,此处的描述对应3-shot(镜头)场景。

[0143] 进一步的,可以使用K最近邻(K-Nearest Neighbors, KNN)算法对样本测试集中的第三样本信道数据集进行初步分类,得到第三样本信道数据对应的第三样本对象标签。之后,基于各第三样本信道数据的评估指标构建样本辅助集。

[0144] 进一步的,可以利用样本训练集的第一样本信道数据,采用交叉熵损失(第一损失)训练神经网络模型。同时,基于样本训练集的目标第一样本信道数据和样本辅助集的第二样本信道数据,在每个样本对象标签中构建MK-MMD损失(第二损失),并将MK-MMD损失与

交叉熵损失函数相结合,构建最终的目标函数对预设模型进行训练。

[0145] 示例性的,可以根据训练出的预设模型在样本支持集上的第二分类准确度来决定何时停止训练。具体而言,当预设模型在样本支持集上的第二分类准确度连续30个训练回合不再提升时,即可停止训练。最终选择样本支持集的第二分类准确度最高的训练轮次对应的第一参数作为最终的模型参数。

[0146] 进一步的,可以直接使用训练好的目标对样本测试集的第三样本信道数据进行分类。实验结果显示,在3-shot (镜头) 场景下,目标模型的分类准确率超过85%,甚至优于在单一环境中使用ResNet网络直接进行分类的性能,这表明本申请能够有效地实现跨环境的行人识别。

[0147] 本申请实施例通过检测第二环境的信道数据的状态;当第二环境的信道数据的状态发生变化时,从第二环境中获取待分析信道数据;将待分析信道数据输入至目标模型中,得到对象标签;其中,目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,目标对象为从多个对象中选取的对象,目标对象的数量少于对象的数量;目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。以此,能够按照同一样本对象标签对齐源域的样本训练集和目标域的样本辅助集,以此,能够通过中间的特征空间使得各样本对象标签在源域和目标域的分布大致相同,以使得模型在少样本训练的情况下,也能够达到良好的训练效果,提高了模型在少样本场景下的训练效率,同时由于目标模型是基于预测结果与样本对象标签之间的差异,以及目标第一样本信道数据与第二样本信道数据的均值差异确定的,由此可以提高训练好的模型对对象识别的准确性。综上,本申请能够在提高模型在少样本场景下的训练效率的同时,提高训练好的模型对对象识别的准确性。

[0148] 请参阅图6,本申请实施例还提供一种基于少样本学习的模型对对象识别的装置,可以实现上述基于少样本学习的模型对对象识别的方法,基于少样本学习的模型对对象识别的装置包括:

检测模块61,用于检测第二环境的信道数据的状态;

获取模块62,用于当第二环境的信道数据的状态发生变化时,从第二环境中获取待分析信道数据;

输入模块63,用于将待分析信道数据输入至目标模型中,得到对象标签;其中,目标模型基于样本训练集中每个第一样本信道数据的第一预测结果与对应的第一样本对象标签之间的差异,以及基于样本训练集中的目标第一样本信道数据与在同一样本对象标签下的样本辅助集中的第二样本信道数据的均值差异确定;样本训练集包括多个对象中针对每个对象在第一环境中的第一样本信道数据以及每个对象的第一样本对象标签;样本辅助集包括目标对象在第二环境中的第二样本信道数据以及每个目标对象的第二样本对象标签,目标对象为从多个对象中选取的对象,目标对象的数量少于对象的数量;目标第一样本信道数据和第二样本信道数据属于同一对象分别在第一环境和第二环境中的信道数据。

[0149] 该基于少样本学习的模型对对象识别的装置的具体实施方式与上述基于少样本学习的模型对对象识别的方法的具体实施例基本相同,在此不再赘述。在满足本申请实施例要求的前提下,基于少样本学习的模型对对象识别的装置还可以设置其他功能模块,以实现上述实施例中的基于少样本学习的模型对对象识别的方法。

[0150] 本申请实施例还提供了一种计算机设备,计算机设备包括存储器和处理器,存储器存储有计算机程序,处理器执行计算机程序时实现上述基于少样本学习的模型对对象识别的方法。该计算机设备可以为包括平板电脑、车载电脑等任意智能终端。

[0151] 请参阅图7,图7示意了另一实施例的计算机设备的硬件结构,计算机设备包括:

处理器71,可以采用通用的CPU(CentralProcessingUnit,中央处理器)、微处理器、应用专用集成电路(ApplicationSpecificIntegratedCircuit,ASIC)、或者一个或多个集成电路等方式实现,用于执行相关程序,以实现本申请实施例所提供的技术方案;

存储器72,可以采用只读存储器(ReadOnlyMemory,ROM)、静态存储设备、动态存储设备或者随机存取存储器(RandomAccessMemory,RAM)等形式实现。存储器72可以存储操作系统和其他应用程序,在通过软件或者固件来实现本说明书实施例所提供的技术方案时,相关的程序代码保存在存储器72中,并由处理器71来调用执行本申请实施例的基于少样本学习的模型对对象识别的方法;

输入/输出接口73,用于实现信息输入及输出;

通信接口74,用于实现本设备与其他设备的通信交互,可以通过有线方式(例如USB、网线等)实现通信,也可以通过无线方式(例如移动网络、WIFI、蓝牙等)实现通信;

总线75,在设备的各个组件(例如处理器71、存储器72、输入/输出接口73和通信接口74)之间传输信息;

其中处理器71、存储器72、输入/输出接口73和通信接口74通过总线75实现彼此之间在设备内部的通信连接。

[0152] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质存储有计算机程序,该计算机程序被处理器执行时实现上述基于少样本学习的模型对对象识别的方法。

[0153] 存储器作为一种非暂态计算机可读存储介质,可用于存储非暂态软件程序以及非暂态性计算机可执行程序。此外,存储器可以包括高速随机存取存储器,还可以包括非暂态存储器,例如至少一个磁盘存储器件、闪存器件、或其他非暂态固态存储器件。在一些实施方式中,存储器可选包括相对于处理器远程设置的存储器,这些远程存储器可以通过网络连接至该处理器。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0154] 本申请实施例描述的实施例是为了更加清楚的说明本申请实施例的技术方案,并不构成对于本申请实施例提供的技术方案的限定,本领域技术人员可知,随着技术的演变和新应用场景的出现,本申请实施例提供的技术方案对于类似的技术问题,同样适用。

[0155] 本领域技术人员可以理解的是,图中示出的技术方案并不构成对本申请实施例的限定,可以包括比图示更多或更少的步骤,或者组合某些步骤,或者不同的步骤。

[0156] 以上所描述的装置实施例仅仅是示意性的,其中作为分离部件说明的单元可以是或者也可以不是物理上分开的,即可以位于一个地方,或者也可以分布到多个网络单元上。

可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0157] 本领域普通技术人员可以理解,上文中所公开方法中的全部或某些步骤、系统、设备中的功能模块/单元可以被实施为软件、固件、硬件及其适当的组合。

[0158] 本申请的说明书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等(如果存在)是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0159] 应当理解,在本申请中,“至少一个(项)”和“若干”是指一个或者多个,“多个”是指两个或两个以上。“和/或”,用于描述关联对象的关联关系,表示可以存在三种关系,例如,“A和/或B”可以表示:只存在A,只存在B以及同时存在A和B三种情况,其中A,B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项(个)”或其类似表达,是指这些项中的任意组合,包括单项(个)或复数项(个)的任意组合。例如,a,b或c中的至少一项(个),可以表示:a,b,c,“a和b”,“a和c”,“b和c”,或“a和b和c”,其中a,b,c可以是单个,也可以是多个。

[0160] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统和方法,可以通过其它的方式实现。例如,以上所描述的系统实施例仅仅是示意性的,例如,上述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0161] 上述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0162] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0163] 集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括多指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本申请各个实施例的方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read-Only Memory,简称ROM)、随机存取存储器(Random Access Memory,简称RAM)、磁碟或者光盘等各种可以存储程序的介质。

[0164] 以上参照附图说明了本申请实施例的优选实施例,并非因此局限本申请实施例的

权利范围。本领域技术人员不脱离本申请实施例的范围和实质内所作的任何修改、等同替换和改进,均应在本申请实施例的权利范围之内。

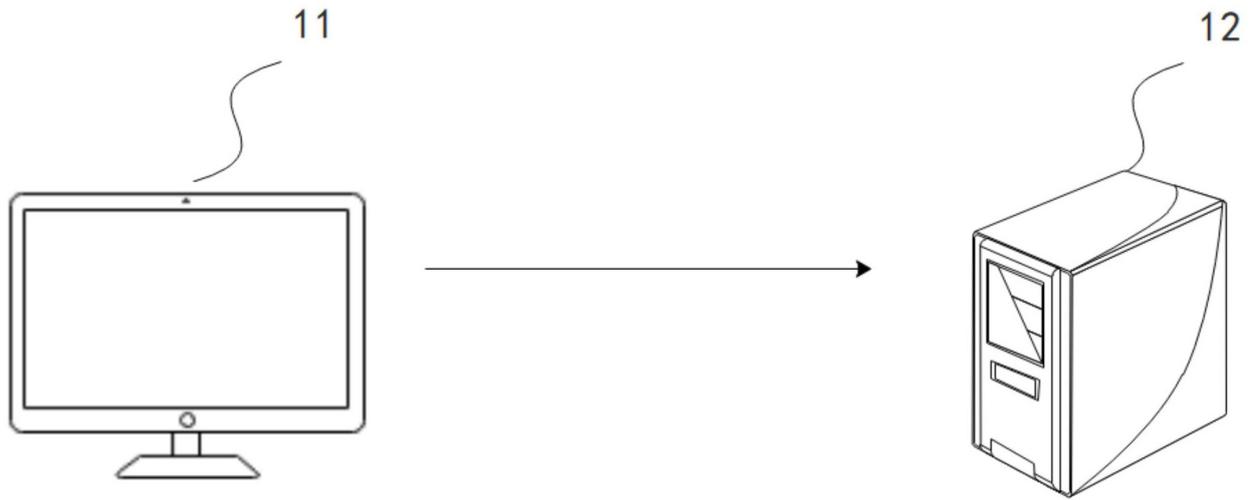


图1

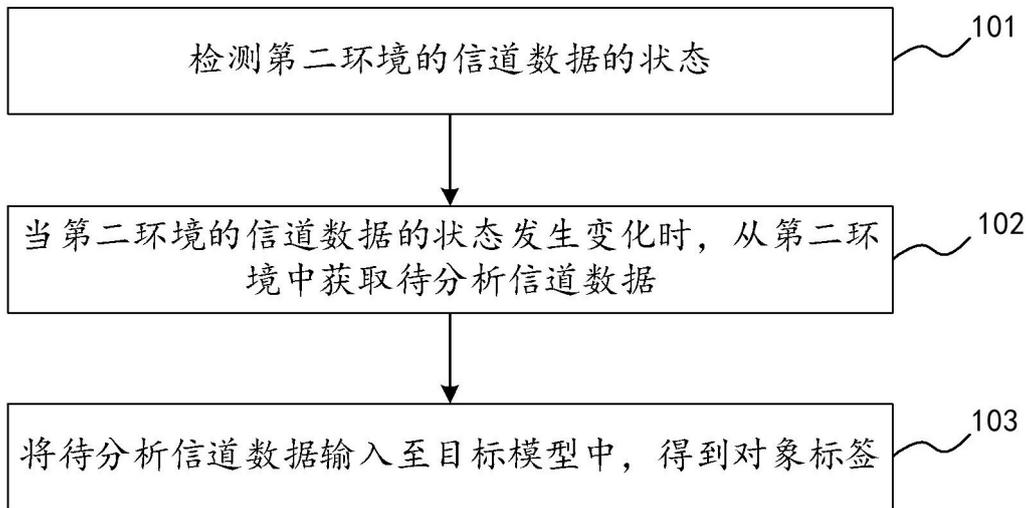


图2

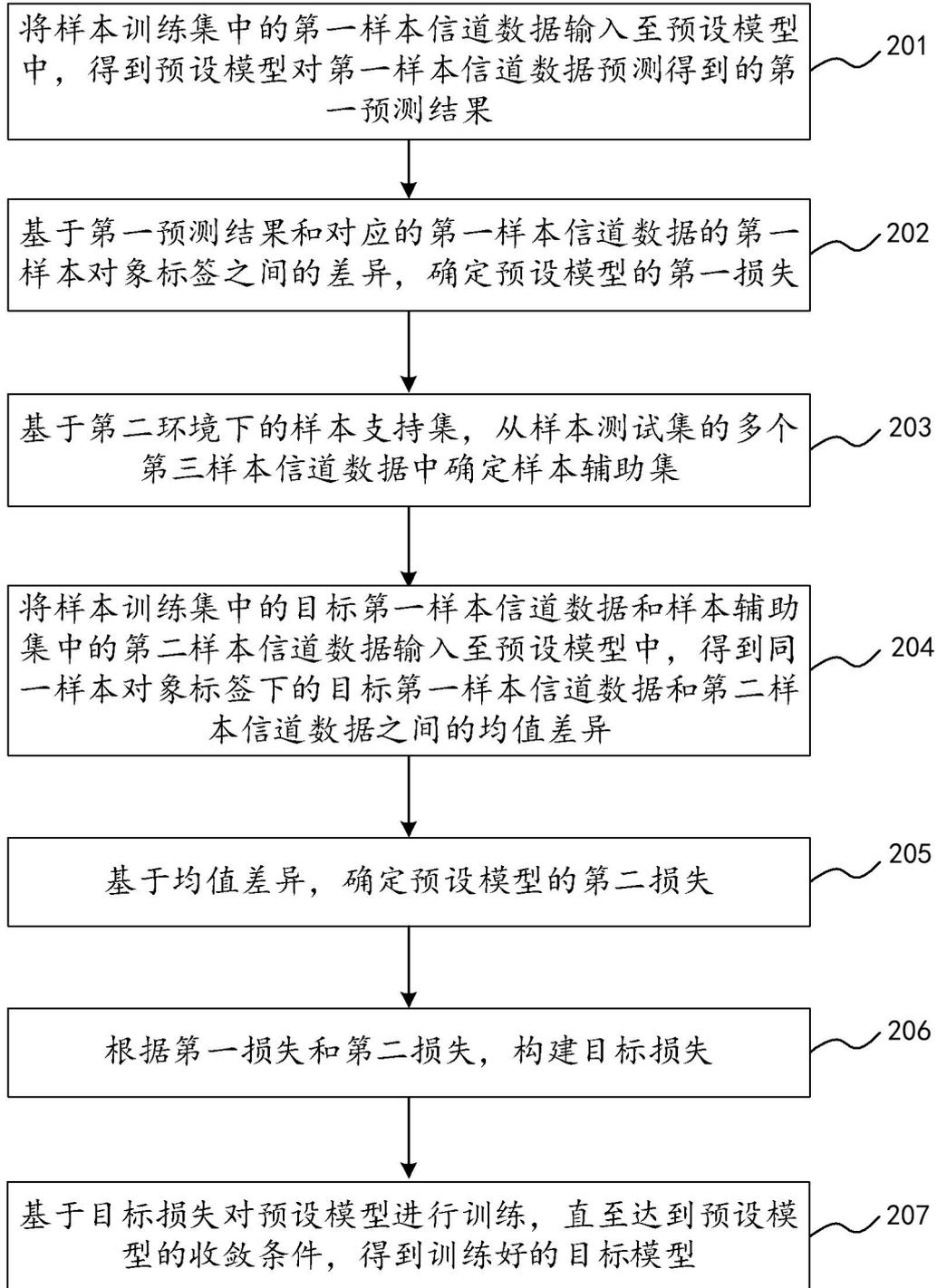


图3

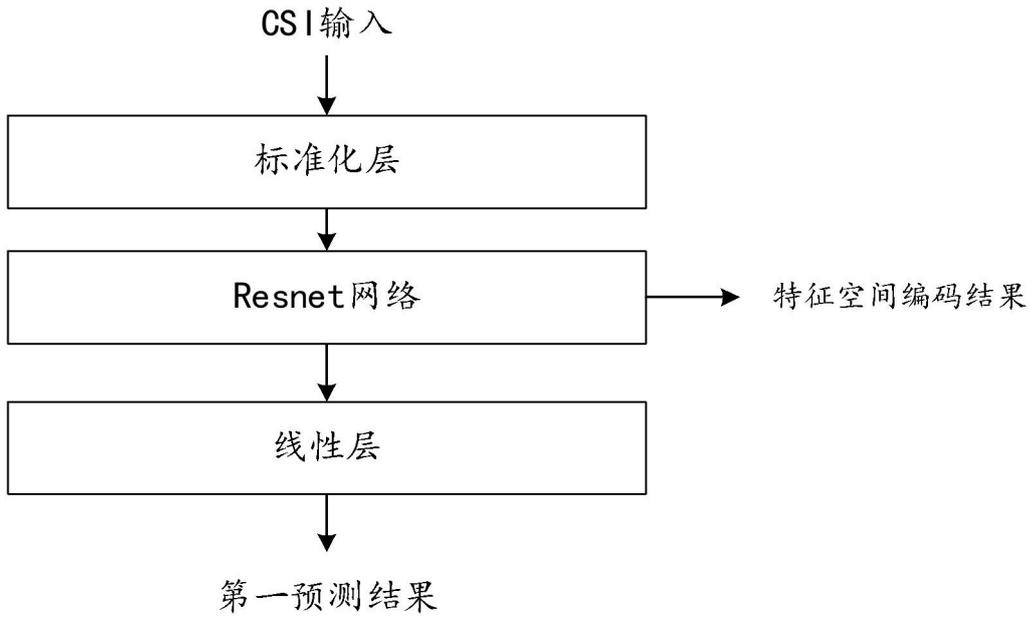


图4

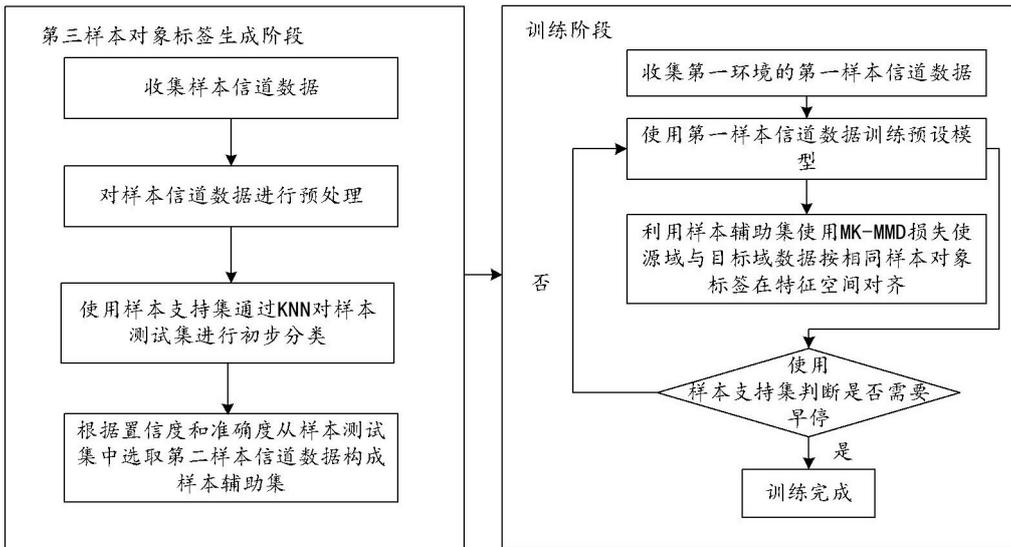


图5

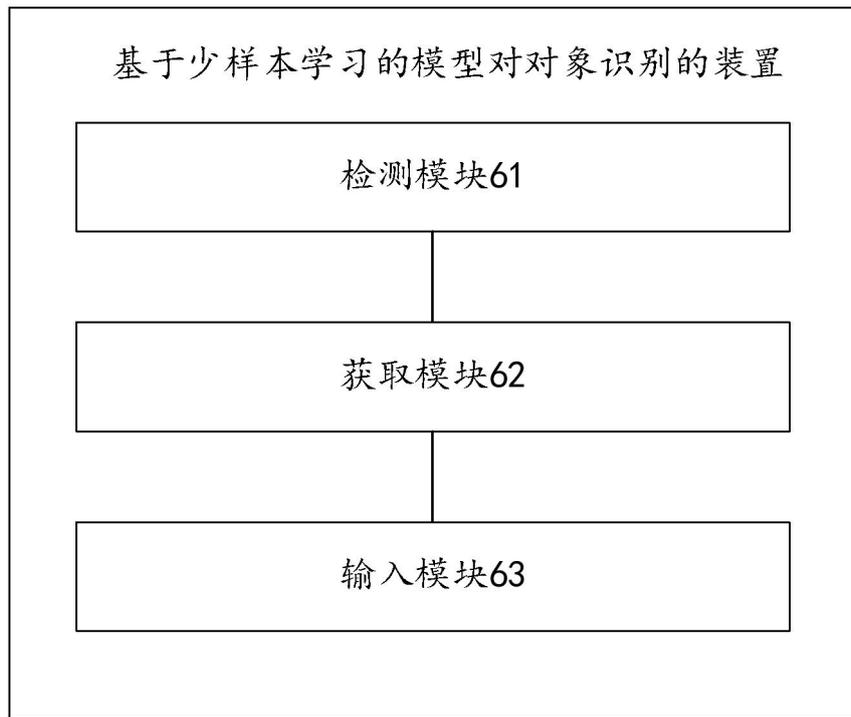


图6

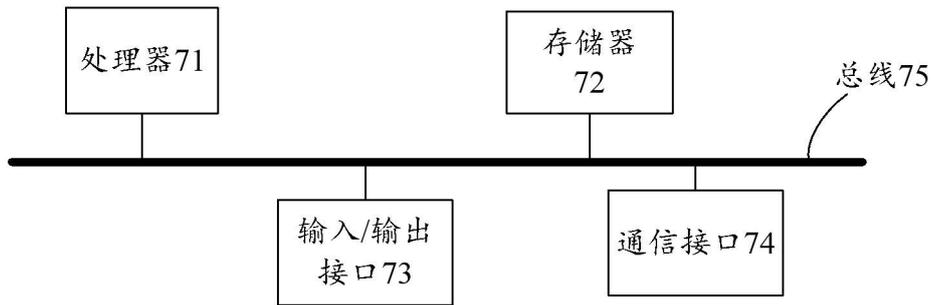


图7