



(12)发明专利申请

(10)申请公布号 CN 108984618 A

(43)申请公布日 2018.12.11

(21)申请号 201810607823.9

(22)申请日 2018.06.13

(71)申请人 深圳市商汤科技有限公司

地址 518000 广东省深圳市南山区南海大道1052号海翔广场712

(72)发明人 吴凌云 罗平 彭章琳 任家敏
张瑞茂 王新江

(74)专利代理机构 北京林达刘知识产权代理事务
所(普通合伙) 11277

代理人 刘新宇

(51)Int.Cl.

G06F 17/30(2006.01)

权利要求书2页 说明书22页 附图9页

(54)发明名称

数据处理方法和装置、电子设备和计算机可读存储介质

(57)摘要

本公开提供了一种数据处理方法和装置、电子设备和计算机可读存储介质,其中所述方法包括:按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据,其中所述多层级标签结构中的第i+1层的标签是由第一层至第i层对应的标签形成的组合标签的扩展标签,其中i为大于或者等于1且小于N的正整数,N为所述多层级标签结构的层数,且N大于1;基于所述多层级标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。本公开实施例能够利用少量的种子标签获取海量标签及对应的数据。

S100

按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据

S200

基于所述多层级标签结构,获取至少一个数据集合

1. 一种数据处理方法,其特征在于,包括:

按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据,其中所述多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签,其中 i 为大于或者等于1且小于 N 的正整数, N 为所述多层级标签结构的层数,且 N 大于1;

基于所述多层级标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。

2. 根据权利要求1所述的方法,其特征在于,所述按照层级扩展的方式,获取基于种子标签的多层级标签结构包括:

将所述种子标签作为与其对应的所述多层级标签结构的第一层标签;

将第 i 层的第一标签以及前 $i-1$ 层与所述第一标签对应的第二标签组合形成组合标签;

获取所述组合标签的扩展标签,所述扩展标签为所述多层级标签结构的第 $i+1$ 层与所述组合标签对应的标签。

3. 根据权利要求2所述的方法,其特征在于,其中,所述获取所述组合标签的扩展标签包括:

通过查询第一数据库,获取所述组合标签的扩展标签,其中所述第一数据库中包括各标签及其扩展标签;或者

通过网络搜索引擎请求获取组合标签的扩展标签。

4. 根据权利要求1所述的方法,其特征在于,所述获取所述多层级标签结构中各标签对应的数据包括:

获得与所述种子标签对应的第一数据;以及

将第 i 层的第一标签以及前 $i-1$ 层与所述第一标签对应的第二标签组合形成组合标签;

获得与所述组合标签对应的第二数据。

5. 根据权利要求4所述的方法,其特征在于,其中,所述获得与所述种子标签对应的第一数据包括:

通过查询第二数据库,获得与所述种子标签对应的第一数据,所述第二数据库中包括多个第一数据,每个第一数据对应有至少一种标签;或者

通过网络搜索引擎获取所述种子标签对应的第一数据。

6. 根据权利要求4所述的方法,其特征在于,其中,所述获得与所述组合标签对应的第二数据包括:

通过查询第二数据库,获得与所述组合标签对应的第二数据,所述第二数据库中包括标签以及与该标签对应的数据;或者

通过网络搜索引擎获取所述组合标签对应的第二数据。

7. 根据权利要求1所述的方法,其特征在于,所述基于所述多层级标签结构,获取至少一个数据集合包括:

利用预设筛选方式,从获得的数据中获取第三数据;

获取第三数据中各数据对应的第三标签;

基于所述第三标签和第三数据形成所述数据集合,

其中,所述预设筛选方式包括随机筛选方式或者基于数据对应的标签的类别执行筛选

的方式。

8. 一种信息处理装置,其特征在于,包括:

扩展模块,其配置为按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据,其中所述多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签,其中 i 为大于或者等于1且小于 N 的正整数, N 为所述多层级标签结构的层数,且 N 大于1;

获取模块,其配置为基于所述多层级标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。

9. 一种电子设备,其特征在于,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为:执行权利要求1至7中任意一项所述的数据处理方法。

10. 一种计算机可读存储介质,其上存储有计算机程序指令,其特征在于,所述计算机程序指令被处理器执行时实现权利要求1至7中任意一项所述的数据处理方法。

数据处理方法和装置、电子设备和计算机可读存储介质

技术领域

[0001] 本公开涉及计算机技术领域,特别涉及一种数据处理方法和装置、电子设备和计算机可读存储介质。

背景技术

[0002] 近年来,深度神经网络在多个智能领域(例如计算机视觉、自然语言处理、游戏博弈等领域)取得了前所未有的技术突破。在视觉理解领域,随着互联网通讯技术的迅猛发展,视觉数据(以图像数据、视频数据为主)的采集成本不断降低,获取并利用海量图像数据(亿级)持续稳定地提升分类神经网络的判别能力与泛化能力成为了可能。然而,海量图像数据分类任务的建模是很复杂的,大规模的数据构建和标签分类都会有很多挑战。

[0003] 现有技术中,大规模数据积累过程通常可以仅使用公开数据集的种子标签作为关键词或使用WordNet(语义网络)语义拓展出的同义词标签作为关键词,再从互联网搜索引擎抓取网络数据,这两种方式都会因其关键词固定使积累数据库的语义概念、样本量及训练模型的迁移性受到限制。其次,学习图像上层信息的深度神经网络建模需要依赖大量高质量、结构化、标签化的数据,但因人工标注成本巨大,在无任何手动数据清洗和人工标注的情况下仍然保证海量数据的结构化和多样化是海量数据构建的难点。

发明内容

[0004] 有鉴于此,本公开提出了一种可以方便的基于少量的种子标签获取海量标签及数据的数据处理方法和装置、电子设备和计算机可读存储介质。

[0005] 根据本公开的第一方面,提供了一种数据处理方法,其包括:

[0006] 按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据,其中所述多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签,其中 i 为大于或者等于1且小于 N 的正整数, N 为所述多层级标签结构的层数,且 N 大于1;

[0007] 基于所述多层级标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。

[0008] 在本公开实施例中,所述按照层级扩展的方式,获取基于种子标签的多层级标签结构包括:

[0009] 将所述种子标签作为与其对应的所述多层级标签结构的第一层标签;

[0010] 将第 i 层的第一标签以及前 $i-1$ 层与所述第一标签对应的第二标签组合形成组合标签;

[0011] 获取所述组合标签的扩展标签,所述扩展标签为所述多层级标签结构的第 $i+1$ 层与所述组合标签对应的标签。

[0012] 在本公开实施例中,所述获取所述组合标签的扩展标签包括:

[0013] 通过查询第一数据库,获取所述组合标签的扩展标签,其中所述第一数据库中包

括各标签及其扩展标签;或者

[0014] 通过向网络搜索引擎请求获取组合标签的扩展标签。

[0015] 在本公开实施例中,所述获取所述多层次标签结构中各标签对应的数据包括:

[0016] 获得与所述种子标签对应的第一数据;以及

[0017] 将第*i*层的第一标签以及前*i-1*层与所述第一标签对应的第二标签组合形成组合标签;

[0018] 获得与所述组合标签对应的第二数据。

[0019] 在本公开实施例中,所述获得与所述种子标签对应的第一数据包括:

[0020] 通过查询第二数据库,获得与所述种子标签对应的第一数据,所述第二数据库中包括多个第一数据,每个第一数据对应至少一种标签;或者

[0021] 利用网络搜索引擎获取所述种子标签对应的第一数据。

[0022] 在本公开实施例中,所述获得与所述组合标签对应的第二数据包括:

[0023] 通过查询第二数据库,获得与所述组合标签对应的第二数据,所述第二数据库中包括标签以及与该标签对应的数据;或者

[0024] 利用网络搜索引擎获取所述组合标签对应的第二数据。

[0025] 在本公开实施例中,所述基于所述多层次标签结构,获取至少一个数据集合包括:

[0026] 利用预设筛选方式,从获得的数据中获取第三数据;

[0027] 获取第三数据中各数据对应的第三标签;

[0028] 基于所述第三标签和第三数据形成所述数据集合,

[0029] 其中,所述预设筛选方式包括随机筛选方式或者基于数据对应的标签的类别执行筛选的方式。

[0030] 在本公开实施例中,所述基于所述多层次标签结构,获取至少一个数据集合包括:

[0031] 对所述多层次标签结构中的各标签进行归类处理,形成关于至少一个类别的标签树结构;

[0032] 基于所述标签树结构获得所述数据集合。

[0033] 在本公开实施例中,所述对所述多层次标签结构中的各标签进行归类处理,形成关于至少一个类别的标签树结构包括:

[0034] 对所述种子标签进行第一归类处理,建立关于至少一个类别的第一标签树结构,所述类别包括物体、地点、事件和时间中的至少一种;

[0035] 对所述种子标签以外的其余标签进行第二归类处理,形成基于所述第一标签树结构的第二标签树结构。

[0036] 在本公开实施例中,所述对所述种子标签进行第一归类处理包括:

[0037] 构建关于各所述种子标签之间的从属关系的第一子标签树结构,其中所述从属关系表示各所述种子标签之间的父类和子类的对应关系;

[0038] 获取各所述第一子标签树结构的根节点的父类信息,如果存在至少两个第一子标签树结构的根节点具有相同的父类,则将该至少两个第一子标签树结构的根节点作为该父类节点的子类进行融合形成新的第一子标签树结构,直至各第一子标签树结构的根节点不存在相同的父类,形成关于至少一个类别的第一标签树结构。

[0039] 在本公开实施例中,所述对所述种子标签以外的其余标签进行第二归类处理,形

成基于所述第一标签树结构的第二标签树结构包括：

[0040] 获取多层次标签结构中第 $i+1$ 层各第四标签的词性；

[0041] 如果所述第四标签的词性为形容词，则将第 $i+1$ 层的该第四标签作为与其对应的前 i 层标签的组合标签的属性信息；

[0042] 如果所述第四标签的词性为名词，则确定该第四标签的对应的前 $i+1$ 层标签的组合标签是否为前 i 层对应的标签的组合标签的子类，如是，将该前 $i+1$ 层标签的组合标签添加至所述第一标签树结构并作为所述前 i 层对应的标签的组合标签的子类，否则，将所述第四标签的词性转化为形容词，并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息；

[0043] 如果所述第四标签的词性为动词，则将第 $i+1$ 层的该第四标签转化为形容词，并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息；

[0044] 如果包括多个第四标签，且该多个第四标签的词性为动词和介词，则该动词和介词词性的第四标签组合作为与其对应的前 i 层标签的组合标签的属性信息。

[0045] 在本公开实施例中，所述基于所述标签树结构获得所述数据集合包括：

[0046] 分别从各类别的所述标签树结构的第 j 层节点中获取一个第五标签，以及与所述第五标签对应的第五数据，其中， j 为大于或者等于1且小于 M 的整数， M 为所述标签树结构的层数；

[0047] 基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0048] 在本公开实施例中，所述基于所述标签树结构获得所述数据集合包括：

[0049] 分别从各所述标签树结构中随机地筛选第五标签，其中，任意两个被筛选的第五标签之间不存在相同的父类；

[0050] 基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0051] 在本公开实施例中，所述方法还包括：

[0052] 分别利用不同的数据集合训练网络模型中的不同分类器；

[0053] 基于满足预设要求的分类器形成所述网络模型。

[0054] 在本公开实施例中，所述分别利用不同的数据集合训练所述网络模型中的不同分类器包括：

[0055] 为每个分类器分配对应的数据集合对所述分类器进行训练，其中各分类器对应的数据集合内的标签不同；或者

[0056] 利用所述数据集合中的第一部分集合训练第一部分分类器，以及利用所述数据集合中的第二部分集合训练第二部分分类器；或者

[0057] 利用所述数据集合中的第一部分集合训练第一部分分类器，以及所述数据集合中的利用第一部分集合和第二部分集合训练第二部分分类器；

[0058] 其中，所述第一部分分类器和第二部分分类器为所述网络模型中的分类器，且所述第一部分分类器和第二部分分类器分别至少包括一个分类器。

[0059] 在本公开实施例中，所述数据包括图像数据、音频数据和视频数据中的至少一种。

[0060] 在本公开实施例中，所述方法还包括通过网络模型预测输入的第三数据的标签类别，其包括：

[0061] 利用各分类器预测第三数据的第一标签类别，以及各第一标签类别的第一预测概

率；

[0062] 基于第一标签树结构,将作为父类节点的第一标签类别的第一预测概率乘以作为其子类的第一标签类别的第一预测概率,得到各作为子类的第一标签类别的第二预测概率；

[0063] 将第二预测概率最高的第一标签类别确定为所述第三数据的标签类别。

[0064] 根据本公开的第二方面提供了一种信息处理装置,其包括：

[0065] 扩展模块,其配置为按照层级扩展的方式,获取基于种子标签的多层级标签结构,以及获取所述多层级标签结构中各标签对应的数据,其中所述多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签,其中 i 为大于或者等于1且小于 N 的正整数, N 为所述多层级标签结构的层数,且 N 大于1；

[0066] 获取模块,其配置为基于所述多层级标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。

[0067] 在本公开实施例中,所述扩展模块进一步配置为将所述种子标签作为与其对应的所述多层级标签结构的第一层标签；

[0068] 将第 i 层的第一标签以及前 $i-1$ 层与所述第一标签对应的第二标签组合形成组合标签；

[0069] 获取所述组合标签的扩展标签,所述扩展标签为所述多层级标签结构的第 $i+1$ 层与所述组合标签对应的标签。

[0070] 在本公开实施例中,所述扩展模块进一步配置为通过查询第一数据库,获取所述组合标签的扩展标签,其中所述第一数据库中包括各标签及其扩展标签；或者

[0071] 通过网络搜索引擎请求获取组合标签的扩展标签。

[0072] 在本公开实施例中,所述扩展模块进一步配置为获得与所述种子标签对应的第一数据；以及将第 i 层的第一标签以及前 $i-1$ 层与所述第一标签对应的第二标签组合形成组合标签,并获得与所述组合标签对应的第二数据。

[0073] 在本公开实施例中,所述扩展模块进一步配置为通过查询第二数据库,获得与所述种子标签对应的第一数据,所述第二数据库中包括多个第一数据,每个第一数据对应至少一种标签；或者

[0074] 利用网络搜索引擎获取所述种子标签对应的第一数据。

[0075] 在本公开实施例中,所述扩展模块进一步配置为通过查询第二数据库,获得与所述组合标签对应的第二数据,所述第二数据库中包括标签以及与该标签对应的数据；或者

[0076] 利用网络搜索引擎获取所述组合标签对应的第二数据。

[0077] 在本公开实施例中,所述获取模块进一步配置为利用预设筛选方式,从获得的数据中获取第三数据,以及获取第三数据中各数据对应的第三标签,并基于所述第三标签和第三数据形成所述数据集合,

[0078] 其中,所述预设筛选方式包括随机筛选方式或者基于数据对应的标签的类别执行筛选的方式。

[0079] 在本公开实施例中,所述获取模块进一步配置为对所述多层级标签结构中的各标签进行归类处理,形成关于至少一个类别的标签树结构,并基于所述标签树结构获得所述数据集合。

[0080] 在本公开实施例中,所述获取模块包括:

[0081] 第一归类单元,其配置为对所述种子标签进行第一归类处理,建立关于至少一个类别的第一标签树结构,所述类别包括物体、地点、事件和时间中的至少一种;

[0082] 第二归类单元,其配置为对所述种子标签以外的其余标签进行第二归类处理,形成基于所述第一标签树结构的第二标签树结构。

[0083] 在本公开实施例中,所述第一归类单元进一步配置为构建关于各所述种子标签之间的从属关系的第一子标签树结构,其中所述从属关系表示各所述种子标签之间的父类和子类的对应关系;

[0084] 并配置为获取各所述第一子标签树结构的根节点的父类信息,如果存在至少两个第一子标签树结构的根节点具有相同的父类,则将至少两个第一子标签树结构的根节点作为该父类节点的子类进行融合形成新的第一子标签树结构,直至各第一子标签树结构的根节点不存在相同的父类,形成关于至少一个类别的第一标签树结构。

[0085] 在本公开实施例中,所述第二归类单元进一步配置为获取多层次标签结构中第 $i+1$ 层各第四标签的词性;

[0086] 如果所述第四标签的词性为形容词,则将第 $i+1$ 层的该第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0087] 如果所述第四标签的词性为名词,则确定该第四标签的对应的前 $i+1$ 层标签的组合标签是否为前 i 层对应的标签的组合标签的子类,如是,将该前 $i+1$ 层标签的组合标签添加至所述第一标签树结构并作为所述前 i 层对应的标签的组合标签的子类,否则,将所述第四标签的词性转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0088] 如果所述第四标签的词性为动词,则将第 $i+1$ 层的该第四标签转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0089] 如果包括多个第四标签,且该多个第四标签的词性为动词和介词,则该动词和介词词性的第四标签组合作为与其对应的前 i 层标签的组合标签的属性信息。

[0090] 在本公开实施例中,所述获取模块进一步配置为分别从各类别的所述标签树结构的第 j 层节点中获取一个第五标签,以及与所述第五标签对应的第五数据,其中, j 为大于或者等于1且小于 M 的整数, M 为所述标签树结构的层数,并基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0091] 在本公开实施例中,所述获取模块进一步配置为分别从各所述标签树结构中随机地筛选第五标签,其中,任意两个被筛选的第五标签之间不存在相同的父类;

[0092] 基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0093] 在本公开实施例中,所述装置还包括:

[0094] 训练模块,其配置为分别利用不同的数据集合训练网络模型中的不同分类器,基于满足预设要求的分类器形成所述网络模型。

[0095] 在本公开实施例中,所述训练模块进一步配置成为每个分类器分配对应的数据集合对所述分类器进行训练,其中各分类器对应的数据集合内的标签不同;或者

[0096] 利用所述数据集合中的第一部分集合训练第一部分分类器,以及利用所述数据集合中的第二部分集合训练第二部分分类器;或者

[0097] 利用所述数据集中的第一部分集合训练第一部分分类器,以及所述数据集中的利用第一部分集合和第二部分集合训练第二部分分类器;

[0098] 其中,所述第一部分分类器和第二部分分类器为所述网络模型中的分类器,且所述第一部分分类器和第二部分分类器分别至少包括一个分类器。

[0099] 在本公开实施例中,所述数据包括图像数据、音频数据和视频数据中的至少一种。

[0100] 在本公开实施例中,所述装置还包括:

[0101] 应用模块,其配置为利用各分类器预测第三数据的第一标签类别,以及各第一标签类别的第一预测概率;

[0102] 基于第一标签树结构,将作为父类节点的第一标签类别的第一预测概率乘以作为其子类的第一标签类别的第一预测概率,得到各作为子类的第一标签类别的第二预测概率;

[0103] 将第二预测概率最高的第一标签类别确定为所述第三数据的标签类别。

[0104] 根据本公开的第三方面提供了一种电子设备,其包括:

[0105] 处理器;

[0106] 用于存储处理器可执行指令的存储器;

[0107] 其中,所述处理器被配置为:执行如上述实施例中任意一项所述的类别数据处理方法。

[0108] 根据本公开的第四方面提供了一种计算机可读存储介质,其上存储有计算机程序指令,其特征在于,所述计算机程序指令被处理器执行时实现上述实施例任意一项所述的数据处理方法。

[0109] 根据本公开实施例的技术方案,能够基于少量的种子标签获得大量的扩展标签数据,并可以建立各标签之间的关联关系,以及根据各标签获取对应的数据信息,从而可以实现利用少量的标签获取众多标签,并可以获取各标签对应的数据信息。

[0110] 根据下面参考附图对示例性实施例的详细说明,本公开的其它特征及方面将变得清楚。

附图说明

[0111] 包含在说明书中并且构成说明书的一部分的附图与说明书一起示出了本公开的示例性实施例、特征和方面,并且用于解释本公开的原理。

[0112] 图1示出根据本公开实施例的一种数据处理方法的流程图;

[0113] 图2示出根据本公开实施例的数据处理方法步骤S100的流程图;

[0114] 图3示出根据本公开实施例的数据处理方法中的多层次标签结构的示意图;

[0115] 图4示出根据本公开实施例的数据处理方法多层次标签结构的示意图;

[0116] 图5示出根据本公开实施例的数据处理方法步骤S100的流程图;

[0117] 图6示出根据本公开实施例的数据处理方法的步骤S200的流程图;

[0118] 图7示出根据本公开实施例数据处理方法形成多层次标签结构以及对应数据的结构的示意图;

[0119] 图8示出根据本公开实施例的数据处理方法获取数据集合的流程图;

[0120] 图9示出根据本公开实施例的数据处理方法形成的标签树结构的结构示意图;

- [0121] 图10示出根据本公开实施例的数据处理方法获得数据集合的流程图；
- [0122] 图11示出根据本公开实施例的数据处理方法获得数据集合的流程图；
- [0123] 图12示出根据本公开实施例的数据处理方法的流程图；
- [0124] 图13示出根据本公开实施例的数据处理方法中步骤S300对分类器进行训练的过程示意图；
- [0125] 图14示出根据本公开实施例的数据处理方法中通过网络模预测输入的第三数据的标签类别的流程图；
- [0126] 图15示出根据本公开实施例的数据处理装置的框图；
- [0127] 图16示出根据本公开一示例性实施例示出的一种电子设备的框图。

具体实施方式

[0128] 以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面，但是除非特别指出，不必按比例绘制附图。

[0129] 在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

[0130] 另外，为了更好的说明本公开，在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解，没有某些具体细节，本公开同样可以实施。在一些实例中，对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述，以便于凸显本公开的主旨。

[0131] 本公开实施例提供了一种数据处理方法，该方法可以用于基于少量的种子标签获得数以百倍千倍甚至更多的扩展标签数据，并可以建立各标签之间的关联关系，以及根据各标签获取对应的数据信息，从而可以实现利用少量的标签获取众多标签，并可以获取各标签对应的数据信息。

[0132] 图1示出根据本公开实施例的一种数据处理方法的流程图。其中，本公开实施例的数据处理方法可以包括：

[0133] S100：按照层级扩展的方式，获取基于种子标签的多层级标签结构，以及获取所述多层级标签结构中各标签对应的数据，其中所述多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签，其中 i 为大于或者等于1且小于 N 的正整数， N 为所述多层级标签结构的层数，且 N 大于1；

[0134] S200：基于所述多层级标签结构，获取至少一个数据集合，所述数据集合包括至少一个数据以及该数据对应的标签。

[0135] 本公开实施例中的数据处理方法可以应用在人工智能或者深度学习领域，例如可以通过该数据处理方法获取用于训练对应的网络模型的数据集合，例如可以利用种子标签以及对、生成的扩展标签获取对应的数据，从而构成上述数据集合。对于不同的网络模型获取的数据可以不同，具体可以根据需求进行设定，例如本公开实施例中的通过标签获得的数据可以为图像数据，在其他实施例中，也可以为音频、视频或者文本等数据，对此，本公开实施例不进行限定。

[0136] 另外，本公开实施例的种子标签是用于获取扩展标签以及数据集合的初始标签，这里的标签相当于索引信息或者关键词信息，具体可以依据不同的需求设定对应的种子标

签,本公开实施例对此不进行限制。

[0137] 本公开实施例中的种子标签包括多个类别的种子标签,以丰富获取的标签以及数据集的内容,例如,上述类别可以包括:物体(object)、场景、时间和事件中的至少一种,或者在其他实施例中也可以设定其他的类别信息,例如,物体可以包括如动物、植物、食材、工具、服饰等,场景可以包括室外、室内、自然地点等,事件可以包括运动、动作、交互关系等,上述仅为列举的实施例,本公开实施例对此不进行限定。

[0138] 另外,本公开实施例中可以从公开的数据集中获取种子标签。其中,公开的数据集可以例如包括ImageNet(图像数据集)、Youtube8M(视频数据集)、Places2(场景图像数据集)、SUN(场景图像数据集)、OpenImages(图像数据集)、ActivityNet(视频分析数据集)、EventNet(事件数据集)、DeepFashion(服装数据集)、Sports1M(运动数据集)、UEC-FOOD(食物数据集)中的至少一种,并且可以通过随机搜索的方式从公开数据集中获取预设数量的种子标签,该预设数量可以根据实际需求进行设定,例如,本公开实施例中的种子标签可以为15个,并且种子标签可以是名词词性的标签。

[0139] 下面在上述说明的基础上对本公开实施例的过程进行详细说明,在步骤S100中,可以按照层级扩展的方式,基于种子标签生成多层级标签结构,以及获取多层级标签结构中各标签对应的数据。本公开实施例中,上述多层级扩展的方式可以将种子标签作为第一层标签,并利用该第一层标签内的种子标签获取下一层标签,以此类推,可以获取多层级标签,基于该多层级标签生成对应的多层级标签结构。其中,本公开实施例中的多层级标签结构中的第 $i+1$ 层的标签是由第一层至第 i 层对应的标签形成的组合标签的扩展标签,其中 i 为大于1且小于 N 的正整数, N 为该多层级标签结构的层数。即第二层标签可以是第一层标签内的种子标签的扩展标签,第三层标签是根据第一层和第二层对应的标签组合形成的组合标签而获得的扩展标签,以此类推,可以获取本公开实施例的多层级标签结构。同时本公开实施例可以利用种子标签获取对应于种子标签对应的数据,也可以利用通过种子标签获得的标签来获取其他层的标签对应的数据,以此类推,可以获取多层级标签结构中各标签对应的数据,并可以建立标签和数据的对应关系。

[0140] 本公开实施例中可以将多层级标签结构设定为 N 层级标签结构,其中, N 为大于1的整数, N 的具体数值可以根据不同的需求进行设定,通常 N 的数值越大,其获取的标签以及对应的数据也就越多,对于后续的数据分析也就越详细,在确定 N 的数值时,也需要综合考虑数据处理能力。

[0141] 图2示出根据本公开实施例中的数据处理方法中步骤S100的流程图。其中步骤S100可以包括:

[0142] S111:将所述种子标签作为与其对应的所述多层级标签结构的第一层标签;

[0143] 即,本公开实施例中的种子标签可以作为对应的多层级标签结构的第一层标签,该种子标签可以为一个,也可以为多个。在种子标签为1个时,可以获取基于该1个种子标签的多层级标签结构,在种子标签为多个时,可以获取基于每个种子标签的多层级标签结构。每个种子标签作为其对应的多层级标签结构的第一层标签。本公开实施例可以针对每个种子标签获取与其对应的多层级标签结构。图3和图4示出了根据本公开实施例中的数据处理方法中多层级标签结构的示意图,其中,以种子标签“蛋糕”为第一层标签,逐渐扩展生成多层扩展标签。

[0144] S112:将第*i*层的第一标签以及前*i*-1层与所述第一标签对应的第二标签组合形成组合标签;

[0145] 如上所述,本公开实施例中的第*i*+1层的标签是由第一层至第*i*层对应的标签形成的组合标签的扩展标签,因此,在获取第2至第*N*层标签时,需要基于之前获取的标签进行组合形成组合标签,即,在获取第*i*+1层标签时,可以将第1至第*i*层对应的第一标签进行组合,形成组合标签。例如,如图3示出,以种子标签为“蛋糕”获取的多层级标签结构中,第二层标签是基于第一层标签组合形成的组合标签获取的,此时第一标签和组合标签都是“蛋糕”。第三层标签是基于第一层和第二层对应的标签组合形成的组合标签获取的,例如,第一标签可以为“城堡”,第二标签即为“蛋糕”,组合标签可以为“城堡蛋糕”,或者第一标签可以为“鲜花”,第二标签即为“蛋糕”,组合标签可以为“鲜花蛋糕”,或者第一标签可以为“生日”,第二标签即为“蛋糕”,组合标签可以为“生日蛋糕”等等。本公开实施例中的第一标签可以为对应层的任意标签,第二标签是与第一标签对应的标签。另外,图3中对于种子标签“奔跑”和“沙滩”的多层级标签结构未一一示出,其用于体现每个种子标签都可以对应的形成其多层级标签结构。

[0146] S113:获取所述组合标签的扩展标签,所述扩展标签为所述多层级标签结构的第*i*+1层与所述组合标签对应的标签。

[0147] 在获取了组合标签之后,可以基于获取该组合标签的扩展标签,并将该扩展标签作为第*i*+1层与上述组合标签对应的标签。

[0148] 本公开实施例中获取扩展标签的方式可以包括:

[0149] 通过查询第一数据库,获取所述组合标签的扩展标签,其中所述第一数据库中包括各标签及其扩展标签对应关系;或者

[0150] 通过向网络搜索引擎请求获取组合标签的扩展标签。

[0151] 其中,第一数据库可以是设置本地设备中的数据库,也可以是设置在云端或者通过网络连接的方式能够获取和查询的数据库,该数据库中存储有不同的词组,以及对应于该词组的扩展词组。本公开实施例中的种子标签可以由词语构成的标签信息,其可以为拼音、中文文字,或者其他语言的拼写等。从而可以通过该第一数据库查找到与种子标签对应的词组,并进一步查找到对应的扩展词组,形成第二层标签,并可以进一步查询到与组合标签对应的扩展词组,形成多层级标签结构。本公开实施例中的第一数据库可以为WordNet数据库,或者其他的能够实现关联标签信息的扩展查询的数据库。

[0152] 或者,在本公开实施例中,也可以通过网络搜索引擎搜索网络中与种子标签或者其他组合标签对应的扩展标签。例如将种子标签作为初始关键词,通过网络搜索引擎(如:谷歌、必应等)获得网络自动扩展的扩展标签作为第二层的标签,同时也可以获取种子标签对应的第一数据。在获取第二层标签后,可以将第一层和第二层中的对应标签组合形成组合标签,并进一步基于网络搜索引擎获得网络自动扩展的扩展标签,作为第三层的标签,以此类推,形成多层级标签结构。

[0153] 通过上述配置可以实现标签的不断扩展,即可以利用少量的种子标签,通过网络搜索引擎或者第一数据库迭代的进行标签扩展和搜索关键词补充。通过这种迭代的标签扩展过程,可以获得数百万扩展标签,远远超出仅使用原始种子标签或使用WordNet扩展语义标签进行弱监督数据积累的方案。

[0154] 另外,在本公开实施例中,在实现标签数据的扩展的同时,还可以实现对应数据的扩展,图5示出根据本公开实施例的数据处理方法中步骤S100的流程图。其中步骤S100可以包括:

[0155] S121:获得与种子标签对应的第一数据;以及

[0156] S122:将所述多层级标签结构的第1层至第i层对应的第一标签组合,形成组合标签;

[0157] S123:获得与所述组合标签对应的第二数据。

[0158] 其中,本公开实施例中,对种子标签以及通过种子标签不断迭代扩展形成的扩展标签分别采用不同的方式获得对应的数据,如上所述,本公开实施例的数据可以根据不同的需求构造为不同形式的数据库,例如该数据可以为图像数据、音频数据、视频数据等等,本公开对此不进行限制。

[0159] 本公开实施例可以直接通过种子标签获得其对应的第一数据。例如,本公开实施例步骤S121可以包括:

[0160] 通过查询第二数据库,获得与所述种子标签对应的第一数据,或者

[0161] 利用网络搜索引擎获取所述种子标签对应的第一数据。

[0162] 本公开实施例中,第二数据库可以是设置在本地设备中的数据库,也可以是设置在云端或者通过网络连接的方式能够获取和查询的数据库,该第二数据库中存储有标签以及与该标签对应的数据,例如,第二数据库中包括多个第一数据,如多个图像,每个第一数据对应至少一种标签。因此,通过种子标签可以在第二数据库中查找到对应的第一数据。

[0163] 或者,在本公开实施例中,也可以通过网络搜索引擎获取对应的第一数据。例如将种子标签作为初始关键词,通过网络搜索引擎(如:Google、bing等)获得搜索到的第一数据。

[0164] 进一步地,本公开实施例还可以通过第1~i层对应标签的组合标签来获得组合标签对应的第二数据。即将所述多层级标签结构的第1层至第i层对应的第一标签组合,形成组合标签。并进一步通过该组合标签获取对应的第二数据。该过程中形成组合标签的过程与上述步骤S112相同,在此不再赘述。另外,在此需要说明的是,本公开实施例中,在获得组合标签对应的数据时,可以不再重复执行标签的组合步骤,即在步骤S112获取组合标签的同时可以执行第二数据的获取,从而节省数据处理时间,避免不必要的功耗。

[0165] 获取第二数据的方式同样可以通过查询第二数据库,来获得与组合标签对应的第二数据,或者利用网络搜索引擎获取所述组合标签对应的第二数据,具体过程与上述实施例相同,在此不再赘述。

[0166] 通过上述配置,即可以获取多层级标签结构以及对应的数据,由于本公开实施例可以通过少量的种子标签获取数以百倍千倍的扩展标签,并可以基于获取的扩展标签进一步获取对应的数据,从而在实现标签的扩展的同时还能实现对应的数据的扩展。图7示出根据本公开实施例数据处理方法形成多层级标签结构以及对应数据的结构的示意图。其中,以种子标签“蛋糕”为例对该过程的展示。第一次,“蛋糕”作为关键词搜索,得到其图像数据和第一层扩展标签“城堡”、“生日”、“鲜花”等(该第一层扩展标签即为多层级标签结构的第二层标签)。第二次,每一个第一层扩展标签与上一次的关键词组合为新的关键词,如“城堡”与“蛋糕”组合为新的关键词“城堡蛋糕”,获得其图像数据和“中世纪的”,“婚礼”,“生

日”等第三层标签。第三次,每个第三层标签与上一次的关键词组合为新的关键词,如“中世纪的”与“城堡蛋糕”组合为新的关键词“中世纪的城堡蛋糕”,获得其图像数据和“简单的,“婚礼”,“中世纪”等第四层标签,依次类推,最终可以获得每个种子标签返回的扩展标签和图像数据,图4为种子标签“蛋糕”的部分扩展标签树结构展示。若“蛋糕”为根节点,按照“蛋糕-城堡-中世纪的-简单的-粉色”分支追溯下去深度为5,树中每一个子节点都可以和其父类节点组合为一个新的搜索关键词,每一个搜索关键词都有对应抓取的图像数据。因此,以这种迭代方式的标签扩展和图像积累,本公开实施例可以获得数百万的扩展标签和数十亿工业级别的数据量,远远超出仅使用种子标签作为搜索关键词进行数据库积累的方式。

[0167] 在获得了多层次标签结构之后,则可以执行基于该多层次标签结构以及对应的数据获取数据集,以完成数据分析或者模型的训练过程。

[0168] 本公开实施例中,通过步骤S100可以实现大量标签以及对应的数据的获取过程,并且标签和数据之间存在对应关系。图6示出根据本公开实施例的数据处理方法的步骤S200的流程图,其中,本公开实施例中的步骤S200可以包括:

[0169] S201:利用预设筛选方式,从获得的数据中获取第三数据;

[0170] S202:获取第三数据中各数据对应的第三标签;

[0171] S203:基于所述第三标签和第三数据形成所述数据集,

[0172] 其中,所述预设筛选方式包括随机筛选方式或者基于数据对应的标签的类别执行筛选的方式。

[0173] 本公开实施例中,每个数据可以至少包括一个对应的标签。因此,在执行数据集的获取时,可以从多层次标签结构中选择预设数量的标签,则可以对应的获取与选择的标签对应的数据,或者可以从获取的数据中选择预设数量的数据,则可以对应的获取与选择的数据对应的标签。通过上述两种方式都可以获得数据集,并且上述预设数量可以根据需求进行设定在此不再赘述。

[0174] 另外,在本公开实施例中,通过上述配置可以获得基于每个种子标签的多层级标签结构,实际应用中各多层次标签结构中的各标签可能存在关联关系,因此本公开实施例可以对各多层次标签进行结构化处理,建立各标签之间的关联,形成基于不同类别的标签树结构,并进一步基于该标签树结构来执行数据集的获取,从而可以建立各标签之间的关联度,并清楚简洁的获取各标签的数据。

[0175] 图8示出根据本公开实施例的数据处理方法获取数据集的流程图,其中步骤S200也可以包括:

[0176] S211:对所述多层次标签结构中的各标签进行归类处理,形成关于至少一个类别的标签树结构;

[0177] S212:基于所述标签树结构获得所述数据集。

[0178] 如上所述,本公开实施例中的种子标签可以包括在不同类别内,如物体、时间、事件、场所等类别。各种子标签之间也可以存在关联,同时形成的各扩展标签与上一层或者前i层标签之间也可以存在关联。本公开实施例可以对各种子标签以及扩展标签与上级标签进行归类分析,建立对应于上述类别中的至少一种的标签树结构,从而可以清晰的建立各标签之间的关联。

[0179] 其中,本公开实施例的步骤S211可以包括:

[0180] 对所述种子标签进行第一归类处理,建立关于至少一个类别的第一标签树结构,所述类别包括物体、地点、事件和时间中的至少一种;

[0181] 对所述种子标签以外的其余标签进行第二归类处理,形成基于所述第一标签树结构的第二标签树结构。

[0182] 如上所述,可以分别对种子标签以及形成的扩展标签进行不同的归类处理,来建立对应类别的标签树结构。例如上述第一归类处理可以包括:

[0183] 构建关于各所述种子标签之间的从属关系的第一子标签树结构,其中所述从属关系表示各所述种子标签之间的父类和子类的对应关系;

[0184] 获取各所述第一子标签树结构的根节点的父类信息,如果存在至少两个第一子标签树结构的根节点具有相同的父类,则可以将该至少两个第一子标签树结构的根节点作为该父类节点的子类进行融合形成新的第一子标签树结构,直至各第一子标签树结构的根节点不存在相同的父类,形成关于至少一个类别的第一标签树结构。

[0185] 本公开实施例中的种子标签可以为名词词性的标签,首先可以通过WordNet数据库或者上述第一数据库,来确定各种子标签之间的从属关系,来建立第一子标签树结构。其中,从属关系表示各所述种子标签之间的父类和子类的对应关系。其中,第一数据库中可以存储有各标签之间的从属关系,即各标签的父类和子类的对应关系。

[0186] 在获取了第一子标签树结构后,可以继续对各第一子标签树结构进行结构化处理,即可以通过第一数据库确定各第一子标签树结构的根节点的父类信息,如果存在至少两个第一子标签树结构的根节点具有相同的父类,这里的父类并不现定于直接的父类,也可以是间隔至少一层的父类,例如可以是两个根节点的父类的父类相同,此时也可以确定为具有相同的父类,则可以将该至少两个第二标签树结构的根节点作为该父类节点的子类进行融合形成新的第一子标签树结构,直至各第一子标签树结构的根节点不存在相同的父类,形成关于至少一个类别的第一标签树结构。

[0187] 也即,本公开实施例中的第一归类处理是基于种子标签之间的从属关系建立至少一个第一子标签树结构,并基于各第一子标签树结构的根节点,向上查找对应的父类,从而对第一子标签树进行结构化整理,形成对应于至少一个类别的第一标签树结构。

[0188] 另外,本公开实施例还可以利用基于种子标签不断迭代获取的扩展标签来补充第一标签树结构,最终形成第二标签树结构。其中,本公开实施例中形成第二标签树结构的过程可以包括:

[0189] 获取所述多层级标签结构第 $i+1$ 层各第四标签的词性;

[0190] 如果所述第四标签的词性为形容词,则将第 i 层的该第四标签作为与其对应的前 $i-1$ 层标签的组合标签的属性信息;例如图2中的“生日蛋糕”的两个扩展标签“粉色的”和“大的”,即可以第一标签树结构中的“生日蛋糕”的属性标签。

[0191] 如果所述第四标签的词性为名词,则确定该第四标签的对应的前 $i+1$ 层标签的组合标签是否为前 i 层对应的标签的组合标签的子类,如是,将该前 $i+1$ 层标签的组合标签添加至所述第一标签树结构并作为所述前 i 层对应的标签的组合标签的子类,否则,将所述第四标签的词性转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;例如,种子标签为“蛋糕”,扩展标签“城堡”,其为名词“城堡”,但它并不是种子标签的子类,所以应该进行形容词化处理,理解为“城堡样子的蛋糕”。至此,可以

将“城堡样子的蛋糕”作为“蛋糕”的子类。

[0192] 如果所述第四标签的词性为动词,则可以将第 $i+1$ 层的该第四标签转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;如“男人”的扩展标签为“奔跑”,此时可以经“奔跑”转换成形容词“正在奔跑的”,并将“正在奔跑的男人”作为“男人”的子类。

[0193] 如果包括多个第四标签,且该多个第四标签的词性为动词和介词,则该动词和介词词性的第四标签组合作为与其对应的前 $i-1$ 层标签的组合标签的属性信息。例如,“沙滩”的扩展标签可以包括“人”、“奔跑”和on,此时可以通过介词将动词节点和单个名词节点(上一层组合标签)相连接,代表该名词正在进行该种动作,即“在沙滩上奔跑的人”。

[0194] 图9示出根据本公开实施例的数据处理方法形成的标签树结构的结构示意图。其中可以获知通过上述配置,即可以实现利用多层次标签结构中除种子标签以外的扩展标签对第一标签树结构进行扩展,向下延伸形成基于第一标签树结构的第二标签树。该第二标签树结构中包括多层次标签结构中的各标签的信息,并包括各标签之间父子类关系,以及对应的属性关系。

[0195] 在获得了第二标签树结构之后,可以进一步基于第二标签树结构选择出数据集合。例如在本公开实施例中,标签树结构的各节点处都至少包括一个对应的标签,以及该标签的属性信息。因此,在执行数据集合的获取时,可以从节点处选择对应的标签,则可以对应的获取与选择的标签对应的数据,进而形成数据集合。

[0196] 另外本公开实施例还可以基于预设规则获取数据集合,图10示出根据本公开实施例的数据处理方法中获得数据集合的流程图。其中,步骤S212可以包括:

[0197] A2121:分别从各类别的所述标签树结构的第 j 层节点中获取一个第五标签,以及与所述第五标签对应的第五数据,其中, j 为大于或者等于1,且小于 M 的整数, M 为所述标签树结构的层数;

[0198] A2122:基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0199] 本公开实施例中,可以形成至少一个类别的标签树结构,该标签树结构的层数可以为 M 层,其中,可以在每个标签树结构的第 j 层获取一个节点处的第五标签,即每个标签树结构的每一层只选择一个标签。这样可以组成互斥关系的标签集合,同时基于各标签对应的数据形成所述数据集合。本公开实施例中互斥关系的标签集合是指该标签集合中的各标签不存在包括相同的直接父类的两个标签。即,每个子类与所属父类不互相排斥,比如:“生日蛋糕”和它的父类“蛋糕”、“食物”、“物体”等都是某张生日蛋糕图像的标签,但它们之间并不互斥。

[0200] 在获取了上述第五标签之后,可以获取与第五标签对应的第五数据,从而基于第五标签和第五数据形成数据集合。

[0201] 或者,图11示出根据本公开实施例的数据处理方法中获得数据集合的流程图。即在本公开另一些实施例中,步骤S220可以包括:

[0202] B2121:分别从各所述标签树结构中随机地筛选第五标签,其中,任意两个被筛选的第五标签之间不存在相同的父类;

[0203] B2122:基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0204] 即本公开实施例,也可以基于形成的标签树结构随机的从其中筛选出多个第五标

签,该被筛选的多个第五标签之间为互斥的关系,即任意两个被筛选的第五标签之间不存在相同的父类。

[0205] 进一步地,可以基于所述第五标签及其对应的第五数据形成至少一个数据集合。

[0206] 通过上述配置,可以实现对多层级标签结构的结构化处理,以及基于结构化处理形成的标签树结构获取对应的数据集合,本公开实施例可以通过上述结构化处理建立标签数据之间的父子类关系,以及对应的属性关系,从而使得标签树结构内的各标签之间的关系更为清楚,继而可以方便的从中获取对应关系的标签及数据。

[0207] 如上述实施例所述,本公开实施例可以应用在人工智能和深度学习领域,本公开实施例可以基于获取的数据集合对网络模型进行训练,以形成具有预设功能的网络模型。而且,本公开实施例可以采用少量的种子标签来形成大量的标签和对应的数据,从而实现对应网络模型的训练,可以大大的提高对网络模型的训练精度。

[0208] 下面对利用上述数据集合训练网络模型的过程进行详细说明。图12示出根据本公开实施例的数据处理方法的流程图。其中除了包括图1中的步骤S100和S200之外,还可以包括:

[0209] S300:分别利用不同的数据集合训练所述网络模型中的不同分类器;

[0210] S400:基于满足预设要求的分类器形成所述网络模型。

[0211] 本公开实施例中,步骤S200可以获取至少一个数据集合,例如,每个类别的标签树结构可以获取一个数据集合,或者可以基于多个类别的标签数据结构内的不同标签可以获取至少一个数据集合,即每个数据集合中可以包括多个类别内的标签,但是每个数据集合中的标签可以不同。

[0212] 即,本公开实施例中步骤S300可以包括:

[0213] 利用特征提取器提取数据集合中的特征信息;

[0214] 为每个分类器分配对应的数据集合的特征信息,以对相应的分类器进行训练,其中各分类器对应的数据集合内的标签不同;或者

[0215] 利用所述数据集合中的第一部分集合的特征信息训练第一部分分类器,以及利用所述数据集合中的第二部分集合的特征信息训练第二部分分类器;或者

[0216] 利用所述数据集合中的第一部分集合的特征信息训练第一部分分类器,以及所述数据集合中的利用第一部分集合的特征信息和第二部分集合的特征信息训练第二部分分类器。

[0217] 其中,第一部分分类器和第二部分分类器为网络模型中的分类器,且第一部分分类器和第二部分分类器均分别至少包括一个分类器,且第一部分集合和第二部分集合共同构成筛选出的数据集合中的至少一部分。

[0218] 在本公开实施例中,可以为每个分类器分配不同的数据集合,即为每个分类器分配的数据集合中的标签是不同的。如利用不同类别对应的第一标签树结构的第一层节点,获取一组标签,该组标签为互斥关系,利用该组标签及其对应的图像数据或者其他数据构成第一数据集合,该第一数据集合可以的特征信息可以用于训练第一分类器。同时利用的第一标签树结构的第二层节点,获取一组标签,利用该组标签及其对应的图像数据或者其他数据构成第二数据集合,该第二数据集合可以的特征信息可以用于训练第二分类器,以此类推,可以为每个分类器分别分配不同的数据集合的特征信息来对分类器进行训练。这

里,第一分类器和第二分类器分别是网络模型中不同的分类器。通过该配置,每层标签对应的数据集可以用于训练一个分类器,从而可以保证分类器内部标签的互斥关系,以及更好提高学习训练精度。在本公开其他实施例中,每个分类器对应的数据集的标签可以不是同一层上的,只要能够对分类器进行训练,即可以作为本公开实施例。

[0219] 另外,在本公开的其他实施例中,网络模型中可以至少存在一个分类器与其他分类器的数据集不同,即数据集中的标签不同,以实现分类器对不同类别的图像的预测。例如,上述实施例所述的将第一部分数据集的特征信息用于训练第一部分分类器,将第二部分数据集特征信息用于训练第二部分分类器。或者。利用第一部分集合的特征信息训练第一部分分类器,以及利用第一部分集合的特征信息和第二部分集合的特征信息训练第二部分分类器。上述方式都可以实现不同分类器的不同类别数据的识别和分类。

[0220] 图13示出根据本公开实施例的数据处理方法中步骤S300对分类器进行训练的过程示意图。

[0221] 其中,通过特征提取器对获取的数据集进行特征提取,获取对应的数据集中数据所对应的特征信息,将对应的数据集的特征信息输入到对应的分类器中进行分类训练。例如,基于物体、场所和事件所对应的标签筛选出的数据集的特征信息可以用于训练第一分类器A,从工具、衣服和食物所对应的标签筛选出的数据集的特征信息可以用于训练第二分类器B,以及从水果、蛋糕、饮料、菜品等标签筛选出的数据集的特征信息可以用于训练第三分类器C等等。即对应不同的分类器可以通过不同的数据集的分类器进行训练,从而通过调节网络参数,如softmax函数,实现网络模型的训练。以上仅为本公开实施例的示例性说明,本公开实施例对此不进行限定。

[0222] 在本公开实施例中,通过上述方式训练各分类器,且各分类器满足预设要求之后,即完成网络模型的训练。基于该网络模型可以执行对应数据的分类任务,即可以向网络模型输入第三数据,经网络模型获得预测结果,输出与第三数据对应的预测标签,以及预测标签的概率。

[0223] 为了提高预测标签的预测精度,本公开实施例可以将标签的层级关系加入网络模型的预测过程。图14示出根据本公开实施例的数据处理方法中通过网络模型预测第三数据的标签类别的流程图。其中可以包括:

[0224] S501:利用各分类器预测第三数据的第一标签类别,以及各第一标签类别的第一预测概率;

[0225] S502:基于第一标签树结构,将作为父类节点的第一标签类别的第一预测概率乘以作为其子类的第一标签类别的第一预测概率,得到各作为子类的第一标签类别的第二预测概率;

[0226] S503:将第二预测概率最高的第一标签类别确定为所述第三数据的标签类别。

[0227] 本公开实施例中,通过建立的网络模型可以对输入的第三数据进行分类识别,其中各分类器可以确定针对第三数据的第一标签类别,以及对应的第一预测概率。本公开实施例中可以基于该第一预测概率进一步提高预测精度。其中可以依据第一标签树结构,将分类器输出的作为父类节点的第一标签类别的第一预测概率乘以作为其子类标签的第一预测概率,得到各子类标签的第二预测概率。其中可以按照条件概率的形式实现每个标签类别的概率预测。比如:如果想计算标签“生日蛋糕”的第一预测概率,需

要其父类标签“蛋糕”的第一预测概率乘以第5层分类器输出的“生日蛋糕”的第一预测概率，而“蛋糕”的第一预测概率又需要其父类标签“食物”的第一预测概率乘以第4层分类器输出的“蛋糕”的低于预测概率，依次类推。用公式的形式表示如下：

[0228] $P_r(\text{birthday cake}) =$

[0229] $P_r(\text{birthday cake} | \text{cake}) * P_r(\text{cake} | \text{food}) * P_r(\text{food} | \text{artifact}) * P_r(\text{artifact} | \text{object}) * P_r(\text{object})$

[0230] 其中， $P_r(\text{object})$ 为第一层分类器直接预测“物体”的概率， $P_r(\text{artifact} | \text{object})$ 为第二层分类器预测的“人工制品”的概率与 $P_r(\text{object})$ 的乘积，即标签为“人工制品”的条件概率，以此类推， $P_r(\text{cake} | \text{food})$ 为标签为“食物”的条件概率， $P_r(\text{birthday cake} | \text{cake})$ 为标签为“生日蛋糕”的条件概率。通过该方式可以获得每个标签的条件概率，最终网络模型输出结果为条件概率最高的子类标签的类别。通过上述方式，可以基于父类标签的预测概率来进一步提高子类标签的预测概率，从而提高了网络模型对于标签类别的预测精度。

[0231] 通过上述配置，可以实现至少一部分分类器的训练数据与其他分类器的训练数据不同，从而可以分别实现不同分类器对于不同数据集合的分类功能。例如可以将后两层标签对应的各个子分类器对应的模型参数进行划分，分布在多个不同的工作节点上，同时将前三层标签对应的子分类器对应的模型参数在每个工作节点上复制一份。每个工作节点负责同一个网络模型的不同部分，而在网络反向传播时，整个模型的全部参数在各个节点上同步更新。或者也可以按照后两层标签在不同工作节点的分布，将整个数据集合划分到不同的工作节点。每个工作节点处理数据集的不同部分，在网络反向传播时，只运用自身的数据产生损失函数计算梯度。各个节点之间共享卷积神经网络的参数，调度节点将汇总每个计算节点的梯度并进行整体模型参数的更新。

[0232] 步骤S400中的预设要求为训练精度达到预设精度，即可以判断为分类器满足预设要求。

[0233] 本公开实施例中，可以利用softmax函数计算分类器的概率分布，因此，可以使用上述实施例中获取的标签为互斥关系的数据集合，对分类器进行训练。

[0234] 综上所述，本公开实施例提供的数据处理方法，可以用于基于少量的种子标签获得数以百倍千倍甚至更多的扩展标签数据，并可以建立各标签之间的关联关系，以及根据各标签获取对应的数据信息，从而可以实现利用少量的标签获取众多标签，并可以获取各标签对应的数据信息。

[0235] 可以理解，本公开提及的上述各个方法实施例，在不违背原理逻辑的情况下，均可以彼此相互结合形成结合后的实施例，限于篇幅，本公开不再赘述。

[0236] 此外，本公开还提供了数据处理装置、电子设备、计算机可读存储介质、程序，上述均可用来实现本公开提供的任一种数据处理方法，相应技术方案和描述和参见方法部分的相应记载，不再赘述。

[0237] 图15示出根据本公开实施例的数据处理装置的框图，其中，该装置可以包括：

[0238] 扩展模块100，其配置为按照层级扩展的方式，获取基于种子标签的多层级标签结构，以及获取所述多层级标签结构中各标签对应的数据，其中所述多层级标签结构中的第*i*+1层的标签是由第一层至第*i*层对应的标签形成的组合标签的扩展标签，其中*i*为大于或者等于1且小于*N*的正整数，*N*为所述多层级标签结构的层数，且*N*大于1；

[0239] 获取模块200,其配置为基于所述多层次标签结构,获取至少一个数据集合,所述数据集合包括至少一个数据以及该数据对应的标签。

[0240] 在本公开实施例中,所述扩展模块进一步配置为将所述种子标签作为与其对应的所述多层次标签结构的第一层标签;

[0241] 将第*i*层的第一标签以及前*i-1*层与所述第一标签对应的第二标签组合形成组合标签;

[0242] 获取所述组合标签的扩展标签,所述扩展标签为所述多层次标签结构的第*i+1*层与所述组合标签对应的标签。

[0243] 在本公开实施例中,所述扩展模块进一步配置为通过查询第一数据库,获取所述组合标签的扩展标签,其中所述第一数据库中包括各标签及其扩展标签;或者

[0244] 通过网络搜索引擎请求获取组合标签的扩展标签。

[0245] 在本公开实施例中,所述扩展模块进一步配置为获得与所述种子标签对应的第一数据;以及将第*i*层的第一标签以及前*i-1*层与所述第一标签对应的第二标签组合形成组合标签,并获得与所述组合标签对应的第二数据。

[0246] 在本公开实施例中,所述扩展模块进一步配置为通过查询第二数据库,获得与所述种子标签对应的第一数据,所述第二数据库中包括多个第一数据,每个第一数据对应至少一种标签;或者

[0247] 利用网络搜索引擎获取所述种子标签对应的第一数据。

[0248] 在本公开实施例中,所述扩展模块进一步配置为通过查询第二数据库,获得与所述组合标签对应的第二数据,所述第二数据库中包括标签以及与该标签对应的数据;或者

[0249] 利用网络搜索引擎获取所述组合标签对应的第二数据。

[0250] 在本公开实施例中,所述获取模块进一步配置为利用预设筛选方式,从获得的数据中获取第三数据,以及获取第三数据中各数据对应的第三标签,并基于所述第三标签和第三数据形成所述数据集合,

[0251] 其中,所述预设筛选方式包括随机筛选方式或者基于数据对应的标签的类别执行筛选的方式。

[0252] 在本公开实施例中,所述获取模块进一步配置为对所述多层次标签结构中的各标签进行归类处理,形成关于至少一个类别的标签树结构,并基于所述标签树结构获得所述数据集合。

[0253] 在本公开实施例中,所述获取模块包括:

[0254] 第一归类单元,其配置为对所述种子标签进行第一归类处理,建立关于至少一个类别的第一标签树结构,所述类别包括物体、地点、事件和时间中的至少一种;

[0255] 第二归类单元,其配置为对所述种子标签以外的其余标签进行第二归类处理,形成基于所述第一标签树结构的第二标签树结构。

[0256] 在本公开实施例中,所述第一归类单元进一步配置为构建关于各所述种子标签之间的从属关系的第一子标签树结构,其中所述从属关系表示各所述种子标签之间的父类和子类的对应关系;

[0257] 并配置为获取各所述第一子标签树结构的根节点的父类信息,如果存在至少两个第一子标签树结构的根节点具有相同的父类,则将至少两个第一子标签树结构的根节点

作为该父类节点的子类进行融合形成新的第一子标签树结构,直至各第一子标签树结构的根节点不存在相同的父类,形成关于至少一个类别的第一标签树结构。

[0258] 在本公开实施例中,所述第二归类单元进一步配置为获取多层次标签结构中第 $i+1$ 层各第四标签的词性;

[0259] 如果所述第四标签的词性为形容词,则将第 $i+1$ 层的该第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0260] 如果所述第四标签的词性为名词,则确定该第四标签的对应的前 $i+1$ 层标签的组合标签是否为前 i 层对应的标签的组合标签的子类,如是,将该前 $i+1$ 层标签的组合标签添加至所述第一标签树结构并作为所述前 i 层对应的标签的组合标签的子类,否则,将所述第四标签的词性转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0261] 如果所述第四标签的词性为动词,则将第 $i+1$ 层的该第四标签转化为形容词,并将转化为形容词的第四标签作为与其对应的前 i 层标签的组合标签的属性信息;

[0262] 如果包括多个第四标签,且该多个第四标签的词性为动词和介词,则该动词和介词词性的第四标签组合作为与其对应的前 i 层标签的组合标签的属性信息。

[0263] 在本公开实施例中,所述获取模块进一步配置为分别从各类别的所述标签树结构的第 j 层节点中获取一个第五标签,以及与所述第五标签对应的第五数据,其中, j 为大于或者等于1且小于 M 的整数, M 为所述标签树结构的层数,并基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0264] 在本公开实施例中,所述获取模块进一步配置为分别从各所述标签树结构中随机地筛选第五标签,其中,任意两个被筛选的第五标签之间不存在相同的父类;

[0265] 基于所述第五标签及其对应的第五数据形成至少一个所述数据集合。

[0266] 在本公开实施例中,所述装置还包括:

[0267] 训练模块,其配置为分别利用不同的数据集合训练网络模型中的不同分类器,基于满足预设要求的分类器形成所述网络模型。

[0268] 在本公开实施例中,所述训练模块进一步配置成为每个分类器分配对应的数据集合对所述分类器进行训练,其中各分类器对应的数据集合内的标签不同;或者

[0269] 利用所述数据集合中的第一部分集合训练第一部分分类器,以及利用所述数据集合中的第二部分集合训练第二部分分类器;或者

[0270] 利用所述数据集合中的第一部分集合训练第一部分分类器,以及所述数据集合中的利用第一部分集合和第二部分集合训练第二部分分类器;

[0271] 其中,所述第一部分分类器和第二部分分类器为所述网络模型中的分类器,且所述第一部分分类器和第二部分分类器分别至少包括一个分类器。

[0272] 在本公开实施例中,所述数据包括图像数据、音频数据和视频数据中的至少一种。

[0273] 在本公开实施例中,所述装置还包括:

[0274] 应用模块,其配置为利用各分类器预测第三数据的第一标签类别,以及各第一标签类别的第一预测概率;

[0275] 基于第一标签树结构,将作为父类节点的第一标签类别的第一预测概率乘以作为其子类的第一标签类别的第一预测概率,得到各作为子类的第一标签类别的第二预测概

率；

[0276] 输出第二预测概率最高的第一标签类别确定为所述第三数据的标签类别。

[0277] 综上所述,本公开实施例提供的数据处理装置,可以用于基于少量的种子标签获得数以百倍千倍甚至更多的扩展标签数据,并可以建立各标签之间的关联关系,以及根据各标签获取对应的数据信息,从而可以实现利用少量的标签获取众多标签,并可以获取各标签对应的数据信息。

[0278] 图16示出根据本公开一示例性实施例示出的一种电子设备的框图。电子设备可以被提供为终端、服务器或其它形态的设备。电子设备可以包括分类数据处理装置800。例如,该装置800可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等终端。

[0279] 参照图16,装置800可以包括以下一个或多个组件:处理组件802,存储器804,电源组件806,多媒体组件808,音频组件810,输入/输出(I/O)的接口812,传感器组件814,以及通信组件816。

[0280] 处理组件802通常控制装置800的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件802可以包括一个或多个模块,便于处理组件802和其他组件之间的交互。例如,处理组件802可以包括多媒体模块,以方便多媒体组件808和处理组件802之间的交互。

[0281] 存储器804被配置为存储各种类型的数据以支持在装置800的操作。这些数据的示例包括用于在装置800上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0282] 电源组件806为装置800的各种组件提供电力。电源组件806可以包括电源管理系统,一个或多个电源,及其他与为装置800生成、管理和分配电力相关联的组件。

[0283] 多媒体组件808包括在所述装置800和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件808包括一个前置摄像头和/或后置摄像头。当装置800处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0284] 音频组件810被配置为输出和/或输入音频信号。例如,音频组件810包括一个麦克风(MIC),当装置800处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中,音频组件810还包括一个扬声器,用于输出音频信号。

[0285] I/O接口812为处理组件802和外围接口模块之间提供接口,上述外围接口模块可

以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0286] 传感器组件814包括一个或多个传感器,用于为装置800提供各个方面的状态评估。例如,传感器组件814可以检测到装置800的打开/关闭状态,组件的相对定位,例如所述组件为装置800的显示器和小键盘,传感器组件814还可以检测装置800或装置800一个组件的位置改变,用户与装置800接触的存在或不存在,装置800方位或加速/减速和装置800的温度变化。传感器组件814可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件814还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0287] 通信组件816被配置为便于装置800和其他设备之间有线或无线方式的通信。装置800可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件816还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0288] 在示例性实施例中,装置800可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0289] 在示例性实施例中,还提供了一种非易失性计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令被处理器执行时实现上述实施例所述的数据处理方法,例如包括计算机程序指令的存储器804,上述计算机程序指令可由装置800的处理器820执行以完成上述方法。

[0290] 本公开可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质,其上载有用于使处理器实现本公开的各个方面的计算机可读程序指令。

[0291] 计算机可读存储介质可以是保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一—但不限于——电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式压缩盘只读存储器(CD-ROM)、数字多功能盘(DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0292] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计

计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0293] 用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构 (ISA) 指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码,所述编程语言包括面向对象的编程语言—诸如 Smalltalk、C++ 等,以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网 (LAN) 或广域网 (WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。在一些实施例中,通过利用计算机可读程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列 (FPGA) 或可编程逻辑阵列 (PLA),该电子电路可以执行计算机可读程序指令,从而实现本公开的各个方面。

[0294] 这里参照根据本公开实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0295] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作,从而,存储有指令的计算机可读介质则包括一个制品,其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0296] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上,使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0297] 附图中的流程图和框图显示了根据本公开的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分,所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0298] 以上已经描述了本公开的各实施例,上述说明是示例性的,并非穷尽性的,并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下,对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中术语的选择,旨

在最好地解释各实施例的原理、实际应用或对市场中的技术的改进,或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

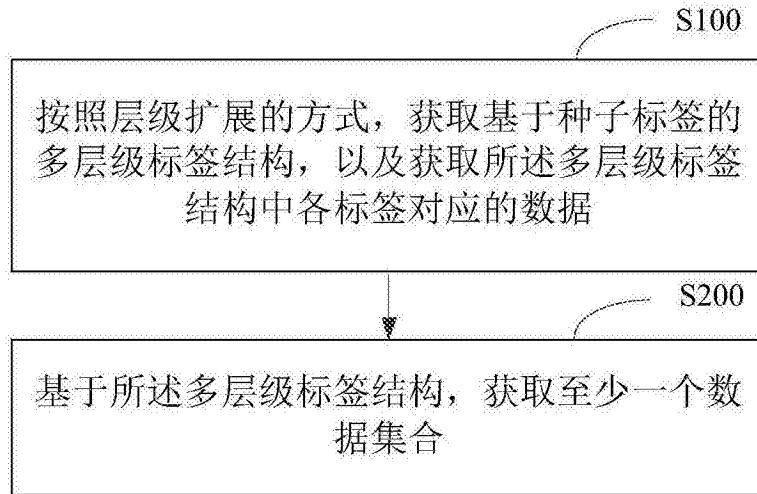


图1

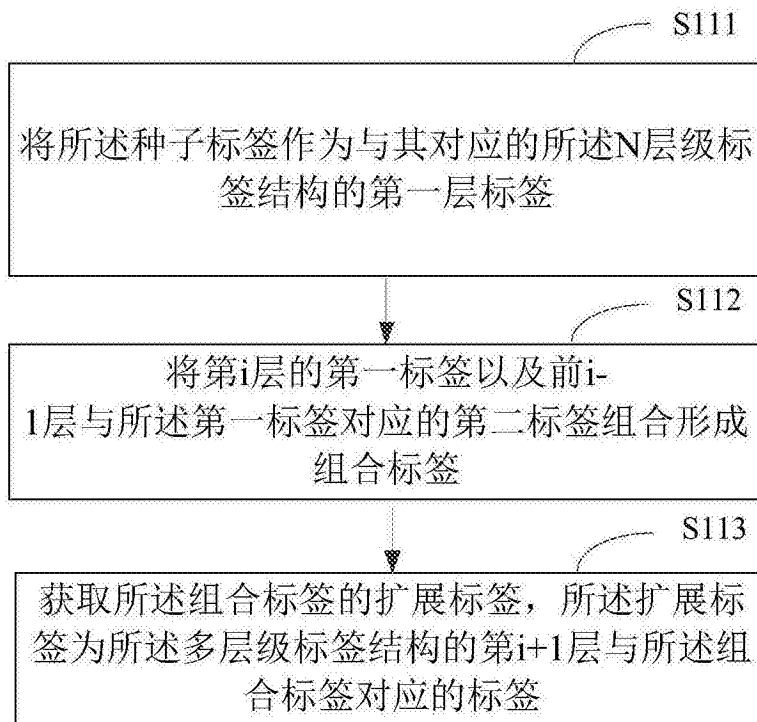


图2

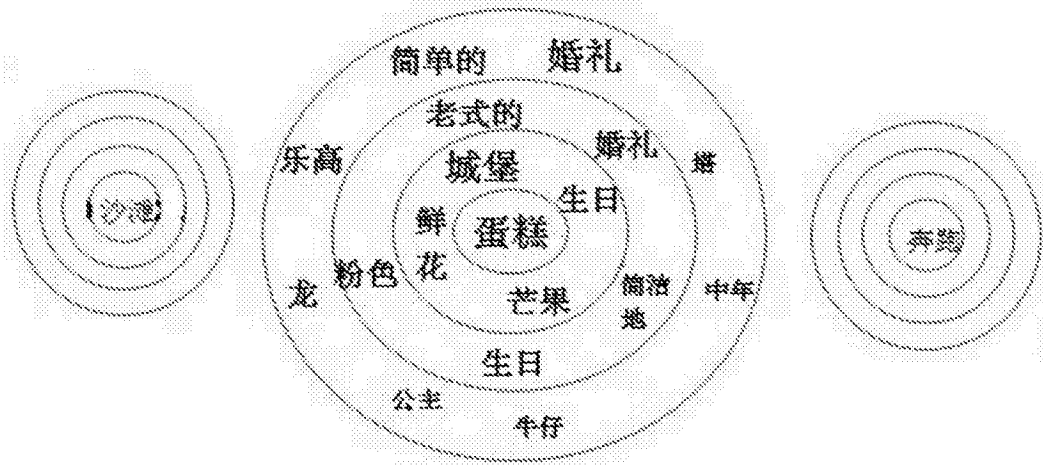


图3

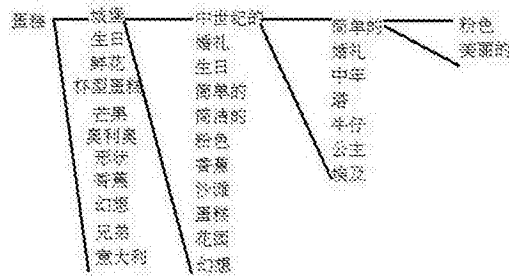


图4

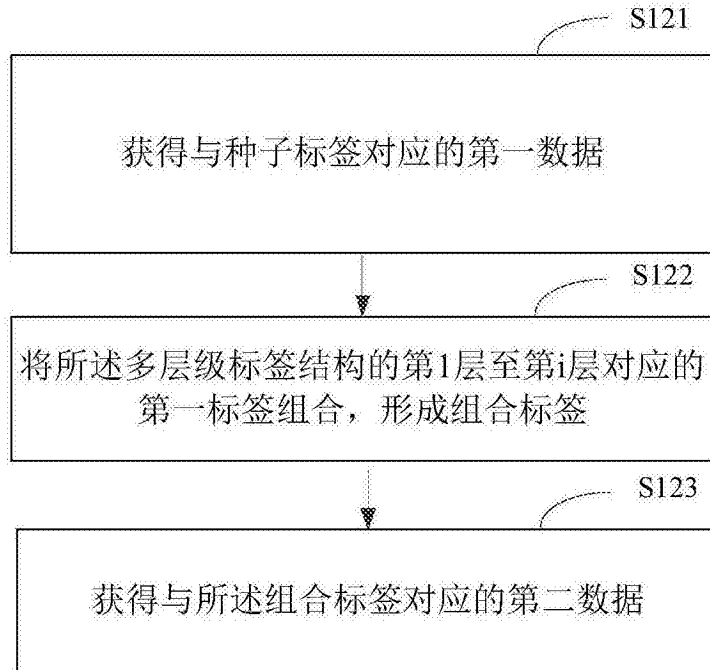


图5

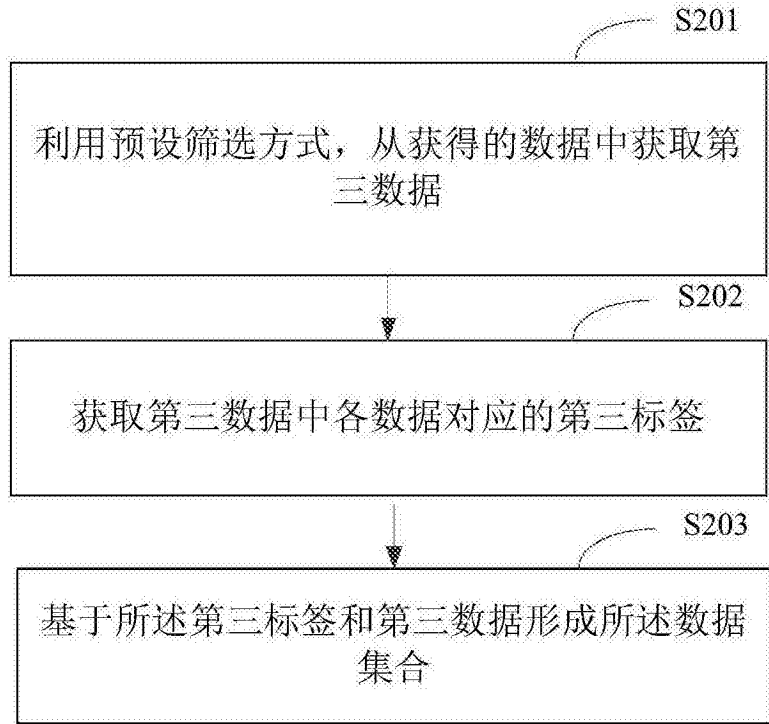


图6

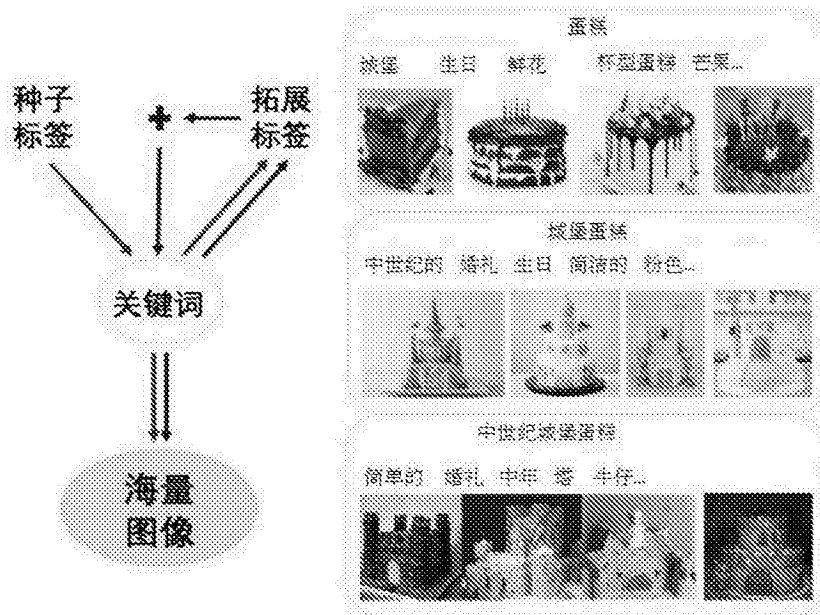


图7

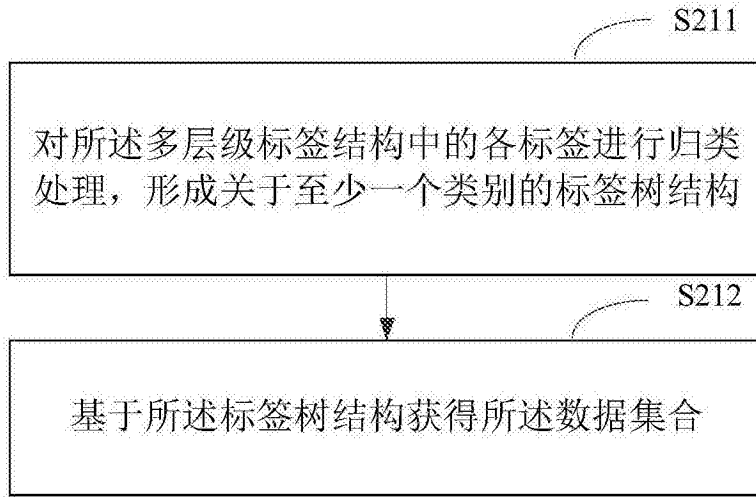


图8

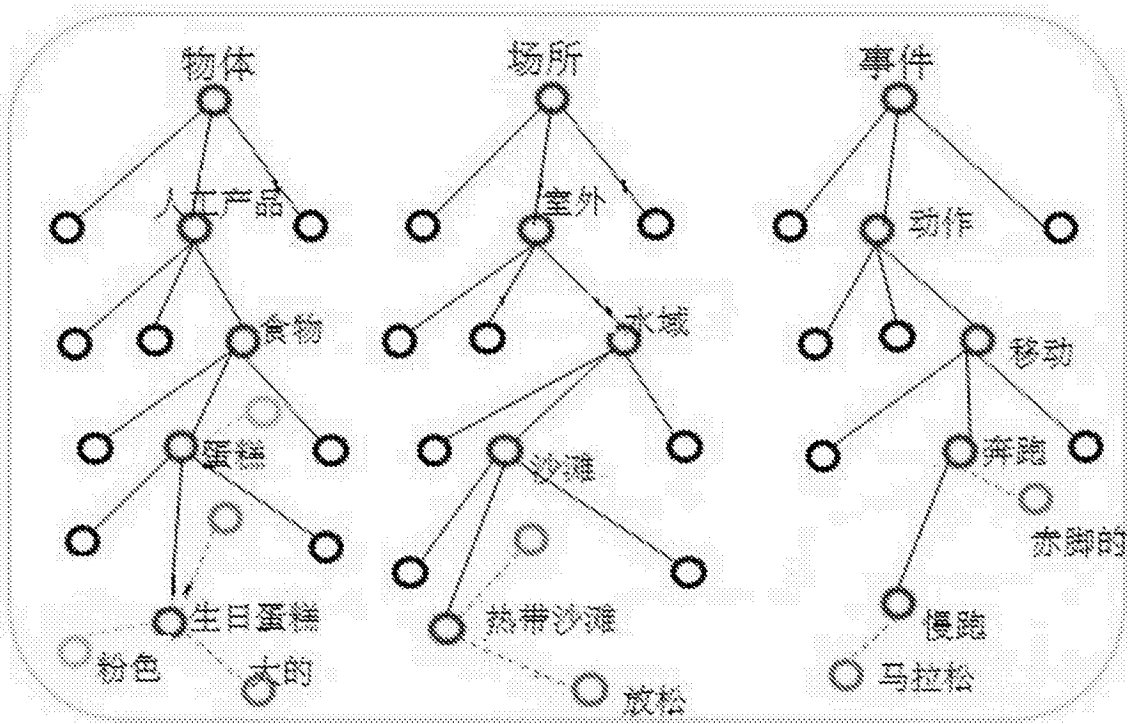


图9

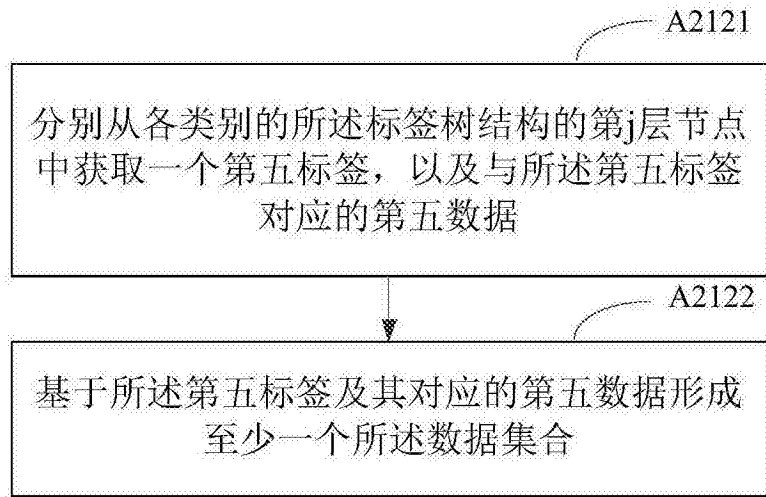


图10

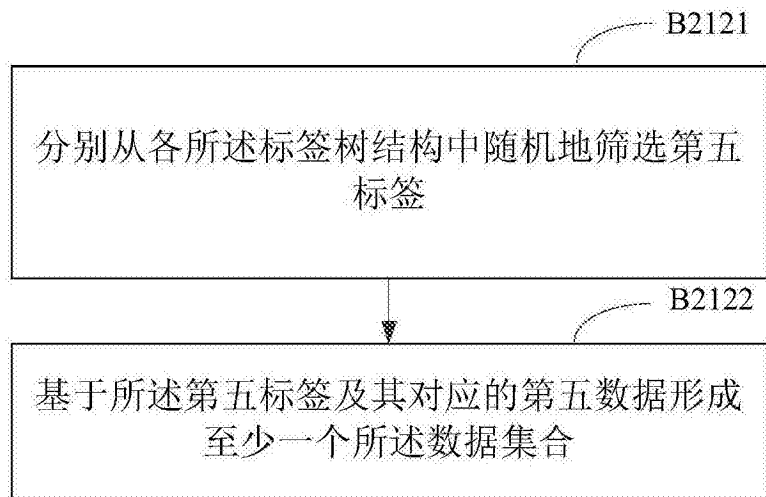


图11

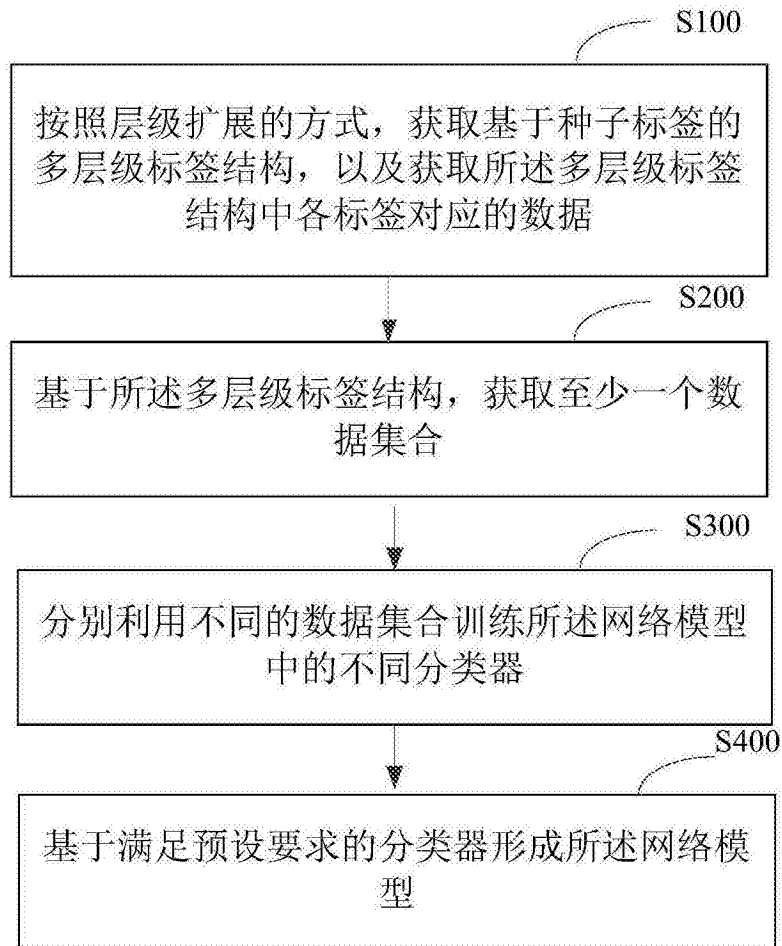


图12

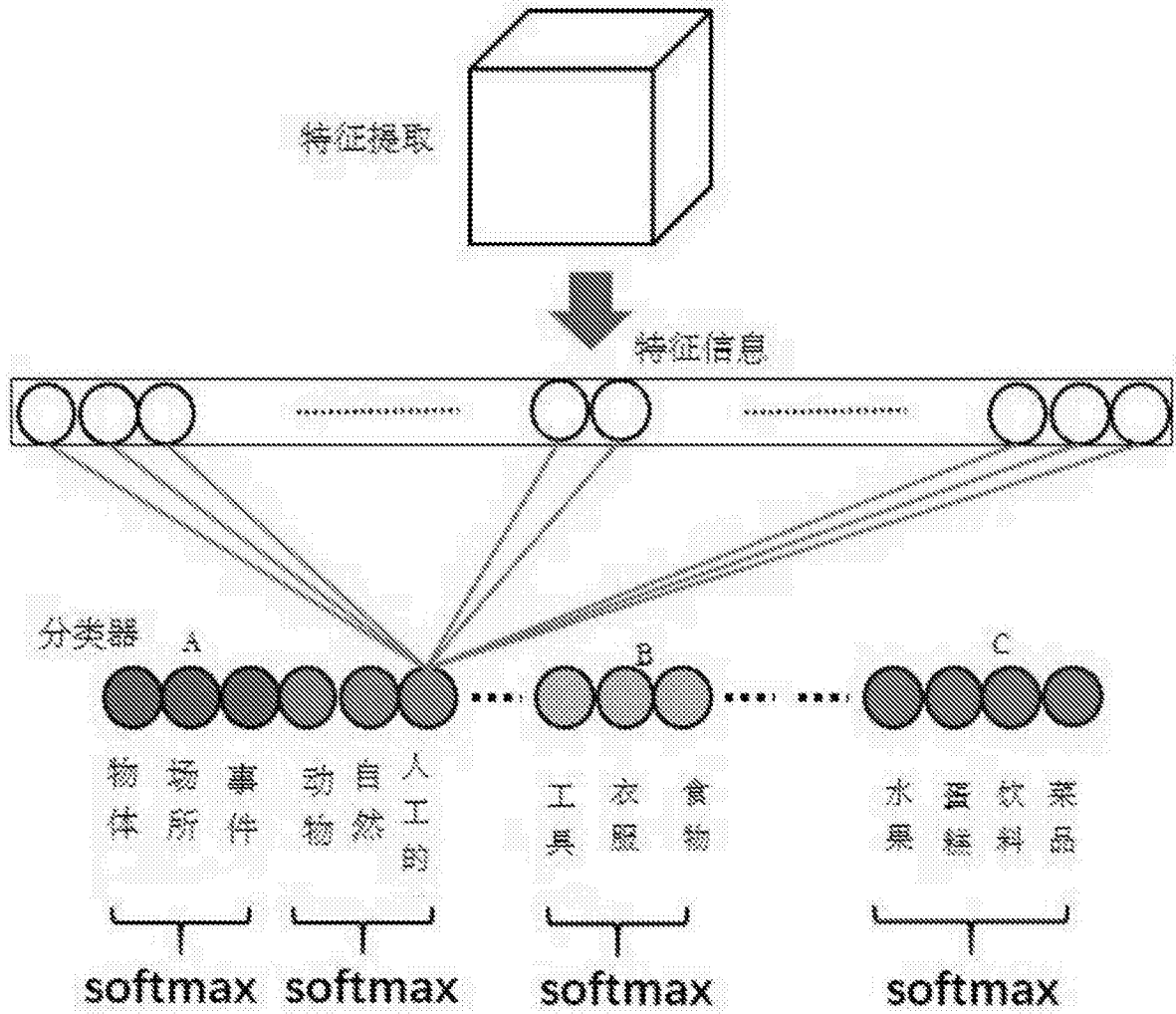


图13

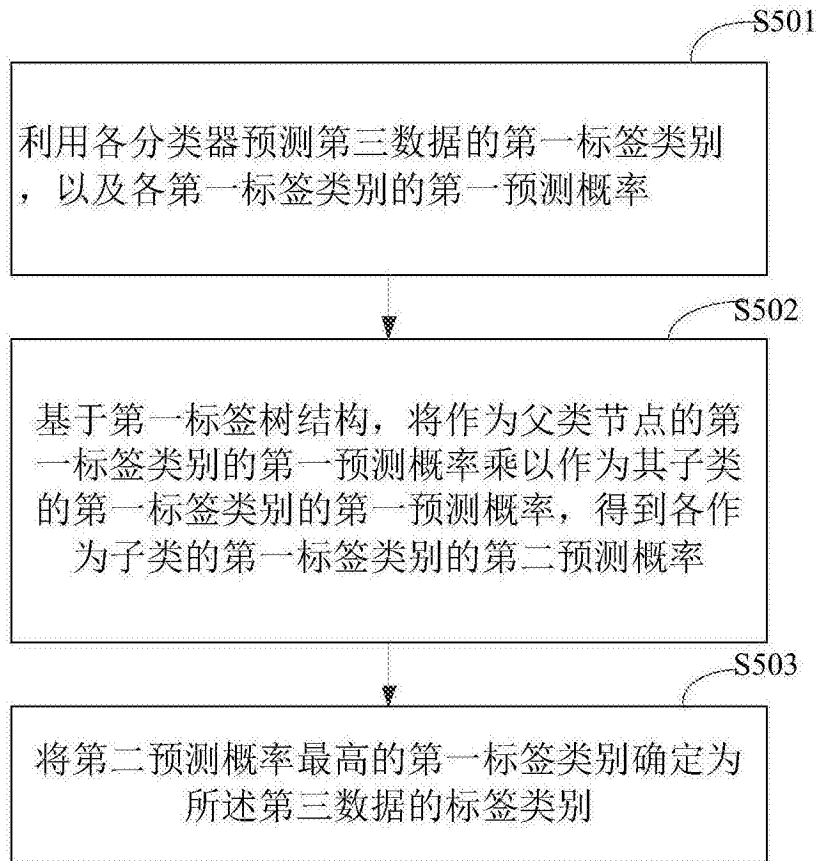


图14

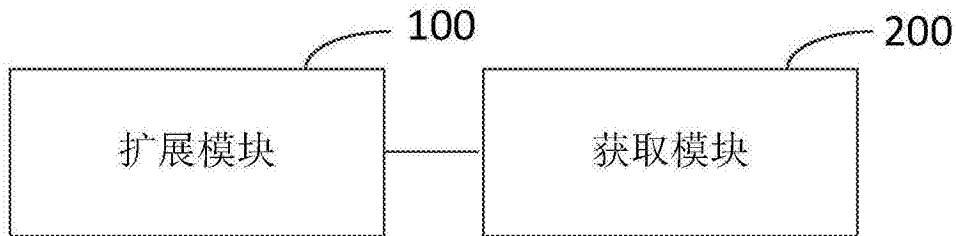


图15

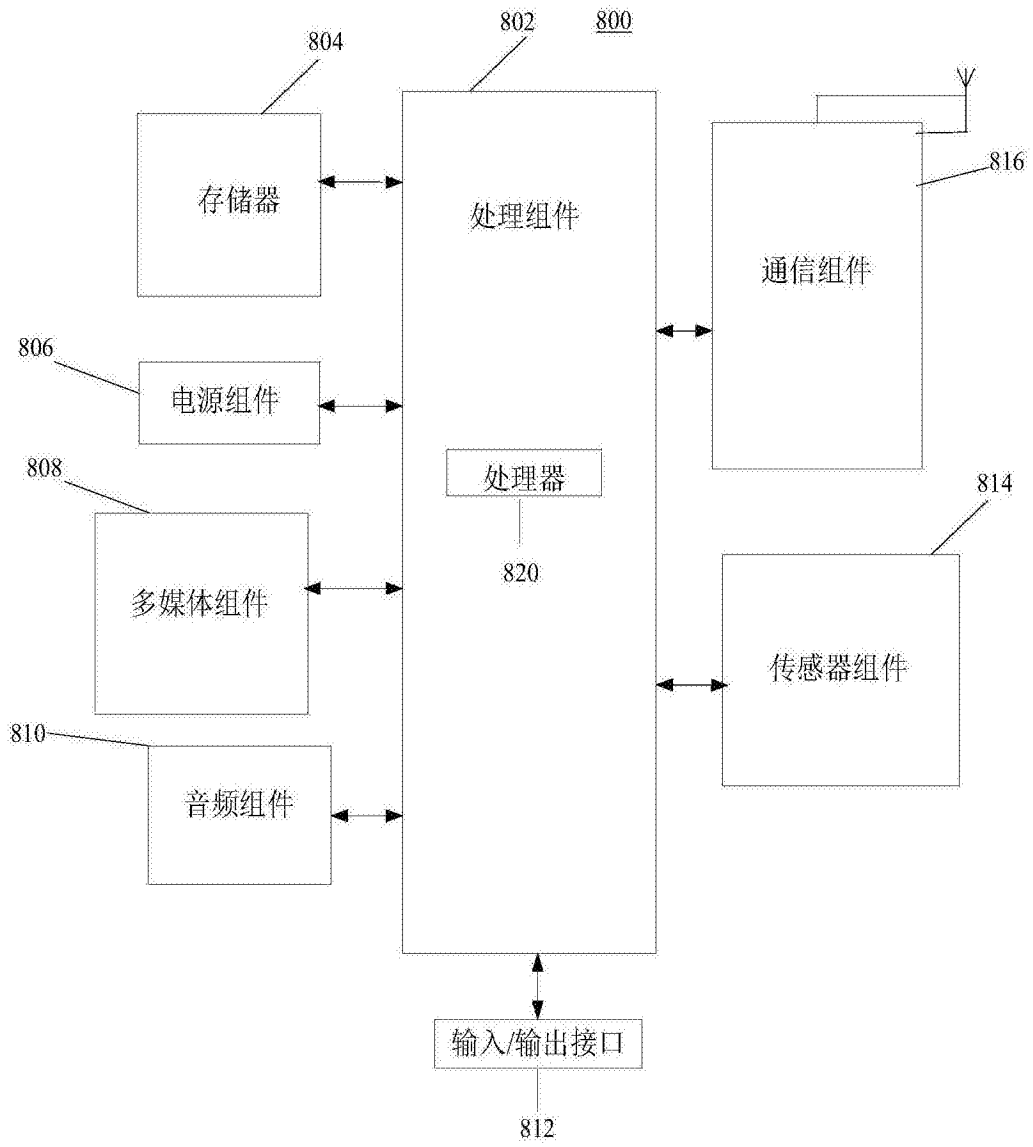


图16