



(12) 发明专利申请

(10) 申请公布号 CN 113989476 A

(43) 申请公布日 2022. 01. 28

(21) 申请号 202111117601.7

(22) 申请日 2021.09.23

(71) 申请人 阿里巴巴(中国)有限公司

地址 310000 浙江省杭州市滨江区长河街
道网商路699号4号楼5楼508室

(72) 发明人 章宦记 孙可嘉 李彤

(74) 专利代理机构 北京众达德权知识产权代理
有限公司 11570

代理人 南海燕

(51) Int. Cl.

G06V 10/22 (2022.01)

G06K 9/62 (2022.01)

G06N 3/08 (2006.01)

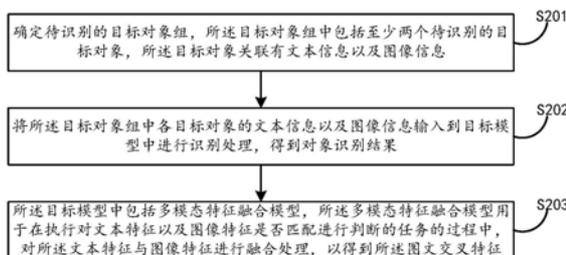
权利要求书3页 说明书17页 附图5页

(54) 发明名称

对象识别方法及电子设备

(57) 摘要

本申请实施例公开了对象识别方法及电子设备,所述方法包括:确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息;将各目标对象的文本信息以及图像信息输入到目标模型中进行识别处理;所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的图文交叉特征,判断各目标对象在目标属性上是否存在共性;所述目标模型中包括多模态特征融合模型,用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征。通过本申请实施例,能够对目标对象在目标属性上是否存在共性进行更准确的识别。



1. 一种对象识别方法,其特征在于,包括:

确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息;

将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行识别处理,得到对象识别结果;其中,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的各目标对象的图文交叉特征,判断各目标对象在目标属性上是否存在共性;

所述目标模型中包括多模态特征融合模型,用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征。

2. 根据权利要求1所述的方法,其特征在于,

所述目标模型包括特征生成模型以及判别模型,所述特征生成模型中包括特征提取模型,以及多模态特征融合模型;

所述特征提取模型用于:分别对各目标对象进行文本特征提取以及图像特征提取;

所述多模态特征融合模型用于:将文本特征与图像特征进行融合,输出各目标对象的图文交叉特征;

所述判别模型用于根据各目标对象的图文交叉特征判断各目标对象在目标属性上是否存在共性。

3. 根据权利要求2所述的方法,其特征在于,

所述目标模型是通过将特征生成模型以及判别模型进行组合而成的成套模型。

4. 根据权利要求3所述的方法,其特征在于,

在对所述目标模型进行训练之前,还包括:

对所述特征生成模型进行训练,并根据训练结果获得所述特征生成模型中的参数学习结果;

将所述参数学习结果作为对所述目标模型进行训练时的参数初始值。

5. 根据权利要求1所述的方法,其特征在于,还包括:

提供关于目标对象在目标属性上是否存在共性的识别结果,并接收关于所述识别结果准确性的用户反馈信息,根据所述用户反馈信息对所述目标模型进行迭代训练。

6. 根据权利要求1至5任一项所述的方法,其特征在于,

所述目标对象包括商品对象;

所述确定各对象在目标属性上是否存在共性,包括:

确定同一商品对象组中的各商品对象是否属于同款或相似款商品对象。

7. 根据权利要求6所述的方法,其特征在于,

所述商品对象组中包括的至少两个商品对象对应相同的类目;

所述目标模型包括对应不同类目的多个不同目标模型;

所述将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,包括:

根据所述目标商品对象组中各商品对象所属的目标类目,将所述各商品对象的文本信息以及图像信息输入到所述目标类目对应的目标模型中进行处理。

8. 根据权利要求6所述的方法,其特征在于,

所述确定待识别的目标对象组,包括:

在根据用户历史浏览过的商品对象进行商品对象推荐时,将该历史浏览过的商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

9. 根据权利要求6所述的方法,其特征在于,

所述确定待识别的目标对象组,包括:

在接收到用户提交的查找目标商品对象的同款/相似款商品对象的请求后,将该目标商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

10. 根据权利要求1至5任一项所述的方法,其特征在于,

所述确定待识别的目标对象组,包括:

接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

所述方法还包括:

根据与所述待识别对象相似度符合条件的目标商品对象对应的标签,提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

11. 一种模型处理方法,其特征在于,包括:

构建图文匹配模型,所述图文匹配模型用于对图像与文本是否用于描述同一对象进行识别;

获取训练样本,所述训练样本中包括多个样本条目,所述样本条目中包括图像内容以及文本内容,以及关于图像内容与文本内容是否匹配的标注信息;

通过所述训练样本对所述图文匹配模型进行训练;

在训练完成后,根据所述图文匹配模型中的特征融合模块,生成所述多模态特征融合模型,以用于对输入的目标对象进行图像特征提取以及文本特征提取,并将图像文本与文本特征进行融合,生成图文交叉特征。

12. 一种提供识别结果的方法,其特征在于,包括:

接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,根据得到的图文交叉特征确定目标对象之间相似度;其中,所述目标模型中包括多模态特征融合模型,所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征;

根据与所述待识别对象的相似度符合条件的目标商品对象对应的标签,提供关于所述

待识别对象是否为正品或与正品的接近程度的识别结果信息。

13. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现权利要求1至12任一项所述的方法的步骤。

14. 一种电子设备,其特征在于,包括:

一个或多个处理器;以及

与所述一个或多个处理器关联的存储器,所述存储器用于存储程序指令,所述程序指令在被所述一个或多个处理器读取执行时,执行权利要求1至12任一项所述的方法的步骤。

对象识别方法及电子设备

技术领域

[0001] 本申请涉及信息识别技术领域,特别是涉及对象识别方法及电子设备。

背景技术

[0002] 在商品对象信息系统中,在很多应用场景中都存在对同款或者相似款商品进行识别的需求。例如,在向用户进行商品推荐时,需要从商品库中识别出与用户历史浏览过的商品相似的商品进行推荐。或者,在一些将多个商品进行聚合展示的页面(例如,活动会场页等等)中,可能会需要将同款或者相似款的商品进行打散之后进行展示,避免同款或相似款的商品过于聚集,此时,也需要从待展示的商品集合中识别出属于同款或者相似款的商品,然后再根据识别结果进行打散处理,等等。

[0003] 在对同款或者相似款商品进行识别时,现有技术中的方案通常是基于商品的图片或者文本进行相似性对比。也即,可以通过将两个商品的图片进行对比,判断两者是否为同款或相似款;或者,通过将两个商品的标题等文本信息进行对比,判断两者是否为同款或相似款。但是图像有可能存在非主图或者主体目标不明显等情况,文本也存在类似属性信息不准确(例如,商家发布商品时,可能存在商品文本信息漏填、错填以及信息堆积冗余,缺乏关键信息等问题)的情况,最终导致同款或相似款的判断结果有偏差。

[0004] 因此,如何提升同款/相似款商品识别的准确率,成为需要本领域技术人员解决的技术问题。

发明内容

[0005] 本申请提供了对象识别方法及电子设备,能够对目标对象在目标属性上是否存在共性进行更准确的识别,在商品对象场景下,可以提升同款/相似款商品识别的准确率。

[0006] 本申请提供了如下方案:

[0007] 一种对象识别方法,包括:

[0008] 确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息;

[0009] 将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行识别处理,得到对象识别结果;其中,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的各目标对象的图文交叉特征,判断各目标对象在目标属性上是否存在共性;

[0010] 所述目标模型中包括多模态特征融合模型,用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征。

[0011] 其中,所述目标模型包括特征生成模型以及判别模型,所述特征生成模型中包括特征提取模型,以及多模态特征融合模型;

[0012] 所述特征提取模型用于:分别对各目标对象进行文本特征提取以及图像特征提

取；

[0013] 所述多模态特征融合模型用于：将文本特征与图像特征进行融合，输出各目标对象的图文交叉特征；

[0014] 所述判别模型用于根据各目标对象的图文交叉特征判断各目标对象在目标属性上是否存在共性。

[0015] 其中，所述目标模型是通过将特征生成模型以及判别模型进行组合而成的成套模型。

[0016] 其中，在对所述目标模型进行训练之前，还包括：

[0017] 对所述特征生成模型进行训练，并根据训练结果获得所述特征生成模型中的参数学习结果；

[0018] 将所述参数学习结果作为对所述目标模型进行训练时的参数初始值。

[0019] 其中，还包括：

[0020] 提供关于目标对象在目标属性上是否存在共性的识别结果，并接收关于所述识别结果准确性的用户反馈信息，根据所述用户反馈信息对所述目标模型进行迭代训练。

[0021] 其中，所述目标对象包括商品对象；

[0022] 所述确定各对象在目标属性上是否存在共性，包括：

[0023] 确定同一商品对象组中的各商品对象是否属于同款或相似款商品对象。

[0024] 其中，所述商品对象组中包括的至少两个商品对象对应相同的类目；

[0025] 所述目标模型包括对应不同类目的多个不同目标模型；

[0026] 所述将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理，包括：

[0027] 根据所述目标商品对象组中各商品对象所属的目标类目，将所述各商品对象的文本信息以及图像信息输入到所述目标类目对应的目标模型中进行处理。

[0028] 其中，所述确定待识别的目标对象组，包括：

[0029] 在根据用户历史浏览过的商品对象进行商品对象推荐时，将该历史浏览过的商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

[0030] 其中，所述确定待识别的目标对象组，包括：

[0031] 在接收到用户提交的查找目标商品对象的同款/相似款商品对象的请求后，将该目标商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

[0032] 其中，所述确定待识别的目标对象组，包括：

[0033] 接收用户提交的待识别对象的文本信息以及图像信息，所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的；

[0034] 将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组；所述商品对象库中的商品对象还关联有标签，所述标签用于表征对应的商品对象是否为正品或与正品的接近程度；

[0035] 所述方法还包括：

[0036] 根据与所述待识别对象相似度符合条件的目标商品对象对应的标签，提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0037] 一种模型处理方法，包括：

[0038] 构建图文匹配模型,所述图文匹配模型用于对图像与文本是否用于描述统一对象进行识别;

[0039] 获取训练样本,所述训练样本中包括多个样本条目,所述样本条目中包括图像内容以及文本内容,以及关于图像内容与文本内容是否匹配的标注信息;

[0040] 通过所述训练样本对所述图文匹配模型进行训练;

[0041] 在训练完成后,根据所述图文匹配模型中的特征融合模块,生成所述多模态特征融合模型,以用于对输入的目标对象进行图像特征提取以及文本特征提取,并将图像文本与文本特征进行融合,生成图文交叉特征。

[0042] 一种模型处理方法,所述模型是由特征生成模型与判别模型进行组合而成的,所述方法包括:

[0043] 对所述特征生成模型进行训练,确定其中的参数值;所述特征生成模型用于将目标对象的图像特征与文本特征进行融合,生成图文交叉特征;

[0044] 获取训练样本,所述训练样本中包括多个目标对象组,所述目标对象组中包括至少两个目标对象,以及关于所述至少两个目标对象是否在目标属性上是否存在共性的标注信息;

[0045] 利用所述训练样本对所述目标模型进行训练,其中,将对所述特征生成模型进行训练过程中确定的参数值,作为对所述目标模型进行训练时的参数的初始值。

[0046] 一种提供识别结果的方法,包括:

[0047] 接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

[0048] 将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

[0049] 将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,根据得到的图文交叉特征确定目标对象之间相似度;其中,所述目标模型中包括多模态特征融合模型,所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征;

[0050] 根据与所述待识别对象的相似度符合条件的目标商品对象对应的标签,提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0051] 一种对象识别装置,包括:

[0052] 目标对象组确定单元,用于确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息;

[0053] 识别处理单元,用于将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行识别处理,得到对象识别结果;其中,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的各目标对象的图文交叉特征,判断各目标对象在目标属性上是否存在共性;

[0054] 所述目标模型中包括多模态特征融合模型,用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所

述图文交叉特征。

[0055] 一种模型处理装置,包括:

[0056] 模型构建单元,用于构建图文匹配模型,所述图文匹配模型用于对图像与文本是否用于描述统一对象进行识别;

[0057] 训练样本获取单元,用于获取训练样本,所述训练样本中包括多个样本条目,所述样本条目中包括图像内容以及文本内容,以及关于图像内容与文本内容是否匹配的标注信息;

[0058] 训练单元,用于通过所述训练样本对所述图文匹配模型进行训练;

[0059] 多模态特征融合模型生成单元,用于在训练完成后,根据所述图文匹配模型中的特征融合模块,生成所述多模态特征融合模型,以用于对输入的目标对象进行图像特征提取以及文本特征提取,并将图像文本与文本特征进行融合,生成图文交叉特征。

[0060] 一种模型处理装置,所述模型是由特征生成模型与判别模型进行组合而成的,所述装置包括:

[0061] 第一训练单元,用于对所述特征生成模型进行训练,确定其中的参数值;所述特征生成模型用于将目标对象的图像特征与文本特征进行融合,生成图文交叉特征;

[0062] 训练样本获取单元,用于获取训练样本,所述训练样本中包括多个目标对象组,每个目标对象组中包括至少两个目标对象,以及关于所述至少两个目标对象是否在目标属性上是否存在共性的标注信息,其中,每个商品对象关联的文本信息以及图像信息;

[0063] 第二训练单元,用于利用所述训练样本对所述目标模型进行训练,其中,将所述对所述特征生成模型进行训练过程中确定的参数值,作为对所述目标模型进行训练时的参数的初始值。

[0064] 一种提供识别结果的装置,包括:

[0065] 信息接收单元,用于接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

[0066] 目标对象组生成单元,用于将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

[0067] 相似度识别单元,用于将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,根据得到的图文交叉特征确定目标对象之间相似度;其中,所述目标模型中包括多模态特征融合模型,所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征;

[0068] 识别结果提供单元,用于根据与所述待识别对象的相似度符合条件的目标商品对象对应的标签,提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0069] 一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现前述任一项所述的方法的步骤。

[0070] 一种电子设备,包括:

[0071] 一个或多个处理器;以及

[0072] 与所述一个或多个处理器关联的存储器,所述存储器用于存储程序指令,所述程序指令在被所述一个或多个处理器读取执行时,执行前述任一项所述的方法的步骤。

[0073] 根据本申请提供的具体实施例,本申请公开了以下技术效果:

[0074] 通过本申请实施例,可以通过将目标对象的多模态的特征(例如,包括文本信息与图像信息等)进行融合的方式,获得目标对象的图文交叉特征,然后可以基于这种图文交叉特征对目标对象在目标属性上是否存在共性进行更准确的识别。其中,在对多模态特征进行融合时,考虑到图像特征与文本特征属于不同域的情况,并不是直接将图像特征与文本特征进行拼接,或者进行简单的数学运算,而是可以构建多模态特征融合模型,并通过有监督学习方式对该模型进行训练,从而可以通过该多模态特征融合模型,对文本特征以及图像特征进行多模态特征的融合。为了实现多模态特征融合模型进行有监督学习方式的训练,还可以将其转化为对图文匹配分类模型的训练,也即,具体的多模态特征融合模型可以是在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,进而可以将融合处理结果确定为图文交叉特征。通过这种方式,可以生成有监督的多模态特征融合模型,以此生成更能表达目标对象特征的图文交叉特征,进而可以更准确地识别出目标对象在目标属性上是否存在共性(例如,各商品对象是否属于同款/相似款)。

[0075] 其中,具体的判别模型也可以是通过有监督学习的方式训练获得的。并且,还可以将包括多模态特征融合模型的特征生成模型与判别模型拼接在一起,组成成套的目标模型,以此实现对特征生成模型与判别模型的同时训练,以提升模型的识别准确度。

[0076] 另外,在对特征生成模型与判别模型组合成的目标模型进行训练之前,还可以首先对特征生成模型进行训练,在确定出该特征生成模型中的参数值之后,可以将其保留下来;之后,可以重新对目标模型进行训练,并且,可以将之前单独训练特征生成模型时确定下来的参数值,作为重新训练目标模型时,模型中的参数的初始值,这样,可以进一步提升模型识别的准确度。

[0077] 当然,实施本申请的任一产品并不一定需要同时达到以上所述的所有优点。

附图说明

[0078] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0079] 图1是本申请实施例提供的系统架构的示意图;

[0080] 图2是本申请实施例提供的第一方法的流程图;

[0081] 图3是本申请实施例提供的多模态融合模型训练过程示意图;

[0082] 图4是本申请实施例提供的另一模型训练过程的示意图;

[0083] 图5是本申请实施例提供的第二方法的流程图;

[0084] 图6是本申请实施例提供的第三方法的流程图;

[0085] 图7是本申请实施例提供的第四方法的流程图;

- [0086] 图8是本申请实施例提供的第一装置的示意图；
[0087] 图9是本申请实施例提供的第二装置的示意图；
[0088] 图10是本申请实施例提供的第三装置的示意图；
[0089] 图11是本申请实施例提供的第四装置的示意图；
[0090] 图12是本申请实施例提供的电子设备的示意图。

具体实施方式

[0091] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员所获得的所有其他实施例,都属于本申请保护的范围。

[0092] 在本申请实施例中,为了提升同款/相似款商品对象识别的准确度,采用了多模态特征相融合的识别方式。具体的,由于商品对象一般可以包括文本(包括标题、品牌名等)以及图像(图片或者视频等)等元素,虽然单纯从文本特征或者图像特征维度进行商品对象识别时,都可能存在识别偏差,但是,如果将文本特征与图像特征融合在一起进行识别,则可以起到相互弥补各自缺点的作用。因此,本申请实施例中,就将文本特征以及图像特征这种单模态的特征融合在一起,生成多模态的图文交叉特征,然后再基于多模态特征进行识别,以此提升同款/相似款商品对象的识别准确度。但是,具体实现时,如何进行文本特征与图像特征的融合,以及具体如何更好地根据融合后的图文交叉特征进行同款/相似款商品对象的识别,是需要考虑的问题。

[0093] 其中,关于多模态特征融合的问题,现有技术中,一种方案是将多种不同模态的特征向量直接拼接在一起,生成多模态特征。例如,图像特征向量是(1,2,3),文本向量是(56,67,90,4),则融合后的特征可以是(1,2,3,56,67,90,4)。或者,还有些方案中,可以在对图像特征以及文本特征进行一些数学运算之后,得到融合后的特征。例如,同样是上述图像特征向量以及文本特征向量,融合后的多模态特征向量可以为(1*56,1*67,1*90,1*4,2*56,2*67,2*90,2*4,3*56,3*67,3*90,3*4),等等。

[0094] 但是,本申请发明人在实现本申请的过程中发现,由于文本特征与图像特征属于不同域的特征,因此,如果直接将两者拼接在一起,或者对两者进行简单的数学运算,融合后的特征可能仍然不足以表达商品对象的特征。为此,在本申请实施例中,首先对文本特征与图像特征的融合方式进行了改进。

[0095] 具体的,为了更合理地进行文本特征与图像特征的融合,可以通过有监督学习的方式建立起多模态特征融合模型,然后通过这种多模态特征融合模型,将同一商品对象的文本特征与图像特征融合为更高级的抽象特征,该抽象特征就可以作为图文交叉特征参与到之后的识别过程中。

[0096] 而在进行有监督学习的模型训练的过程中,由于需要预先对样本进行标注(通常可以是人工标注的方式),标注之前需要先知道训练的目标(也就是期望模型输出的信息)。而如果一个模型直接的输出结果是特征向量,是难以进行样本的人工标注的(也正是因为此原因,现有技术中类似的特征生成模型都是通过无监督的方式进行训练的)。为此,在本申请实施例中,首先可以构建出用于执行对商品对象的文本特征以及图像特征是否匹配

(也即,判断输入的一条文本与一条图像是否可用于描述同一商品对象)进行判断任务的模型。由于在对模型进行训练使其能够判断出文本特征以及图像特征是否匹配的过程中,主要需要学习的就是如何将文本特征与图像特性融合为更高级的特征向量,也就是说,在执行对文本特征以及图像特征是否匹配进行判断的任务过程中,其中一层的输出结果就会包括对文本特征与图像特征进行融合后的特征向量,因此,就可以将该向量作为本申请实施例中融合后的图文交叉向量,以用于参与到后续的同款/相似款商品对象的识别过程中。

[0097] 换言之,本申请实施例中,将对多模态特征融合模型的训练,转化为对图文匹配分类模型的训练。这样可以构建出用于进行图文匹配分类的模型,例如,具体可以基于Bert模型进行构建,等等。然后,可以通过有监督的方式对图文匹配分类模型进行训练。具体的,可以将多个<文本,图像>对以及对应的标注信息作为训练样本,其中,标注信息就可以是文本与图像是否匹配。完成训练后,该模型的作用就是,针对输入的文本与图像,给出两者是否匹配的判断结果。当然,由于本申请实施例中实际上并不需要对图像与文本的匹配情况进行分类,而只需要获取文本特征与图像特征的融合结果,因此,可以从上述训练出的模型中,将从模型中将最终用于输出图文匹配判断结果的一层(也即,输出层)去掉,这样,在输入商品对象的文本以及图像之后,就可以输入将文本特征与图像特征进行融合后的图文交叉特征向量。

[0098] 通过上述方式,可以实现基于有监督学习的方式,将文本特征与图像特征进行融合,而不是简单的进行两个特征向量的拼接,因此,可以实现对商品对象更高级的抽象表达,这种融合处理得到的图文交叉特征能够更全面更准确地表达商品对象的特征。

[0099] 在得到商品对象的图文交叉特征之后,具体在基于这种图文交叉特征进行同款/相似款商品对象的识别时,可以有多种方式。在本申请的优选实施例中,具体的判别模型也可以是通过有监督学习的方式训练获得的。并且,还可以将特征生成模型与判别模型拼接在一起,组成目标模型,以此实现对特征生成模型与判别模型的同时训练。

[0100] 另外,为了进一步提升识别的准确度,在对特征生成模型与判别模型组合成的目标模型进行训练之前,还可以首先对特征生成模型进行训练,在确定出该特征生成模型中的参数值之后,可以将其保留下来;之后,可以重新对目标模型进行训练,并且,可以将之前单独训练特征生成模型时确定下来的参数值,作为重新训练目标模型时,模型中的参数的初始值,这样,可以进一步提升模型识别的准确度。

[0101] 从系统架构角度而言,本申请实施例提供了对目标对象是否在目标属性上存在共性进行识别的功能,该功能可以作为独立的应用程序提供给用户使用,或者,还可以作为其他系统的功能模块而存在。例如,如图1所示,可以在商品对象信息系统中提供该对象识别功能模块,这样,当商品对象信息系统中的其他功能模块需要判断某些商品对象是否属于同款或者相似款时,就可以通过该对象识别模型识别。例如,当推荐模块需要向用户推荐与其历史浏览商品相似的商品对象时,就可以利用该功能模块从商品池中进行同款/相似款商品对象的获取。或者,需要将某页面中的商品对象按款式进行打散时,也可以利用该功能模块进行同款/相似款商品对象的识别后进行打散处理。或者,有的商品对象信息系统还可能为用户提供“找相似”或者“找同款”等功能模块,例如,在某商品对象列表页中,通过多个资源位展示出多个商品对象的过程中,还可以在具体的资源位中提供“找相似”等操作控件,在用户点击该操作控件后,可以将该商品对象与商品库中的商品对象进行同款/相似款

识别,并返回识别结果,等等。此时,也可以由上述“找相似”或者“找同款”等功能模块,确定出具体需要识别的商品对象组,并通过调用上述对象识别功能模块,对组内的各商品对象是否属于同款或者相似款进行识别,等等。

[0102] 或者,本申请实施例还可以利用具体的同款/相似款商品对象识别能力,为用户提供其他更新形式的服务。例如,在实际应用中,某用户可能通过某渠道购买了一个商品,可能需要判断该商品是否为正品;或者,某用户A与用户B分别购买了同品牌的同款化妆品,此时,可能需要对这两个商品进行对比,看哪个更像正品,等等。因此,可以为用户提供上述服务,具体的,可以在商品对象信息系统中提供相关的功能模块(或者轻应用、小程序等),或者提供独立的应用程序,或者,可以提供手持的硬件设备,等等。通过该服务,用户可以将其购买到的商品(对应有实物)作为待识别对象,并通过对其拍照等方式,进行图像信息的采集,并且可以上传具体的图像信息。该服务则可以基于该图像信息,与商品库中各相关商品进行相似性识别,另外,商品库中的商品可以预先添加有是否为正品的标签,或者与正品接近程度的标签,等等。这样,如果识别出当前待识别对象与商品库中某个商品对象的相似度比较高,则可以利用该商品对象对应的标签,返回关于所述待识别对象是否为正品或与正品的接近程度的识别结果。例如,“您的商品可能是正品”等。或者,在上传了两个待识别对象的照片进行正品率对比的情况下,可以分别将两个待识别对象与商品库中的各相关商品进行相似度识别,然后,根据与各自相似度比较高的商品对应的标签,给出两个待识别对象哪个更像正品等识别结果,等等。

[0103] 这里需要说明的是,在根据用户上传的图像信息等与商品库中的商品对象进行同款或者相似款识别时,用户上传的图像信息中可能不存在文本内容,此时,可以直接将图像信息中的图像特征进行更高级的特征抽象。或者,还可以从图片中识别出文本,例如,包括商标文本等等,进而进行文本特征的抽取,又或者,还可以为用户提供用于输入具体文本内容的输入控件,使得用户可以为具体的待识别对象输入文本内容,以提升识别准确度,等等。

[0104] 另外需要说明的是,本申请实施例提供的方案也可以不限于在同款/相似款商品对象识别的场景中使用,而是也可以扩展到其他领域。例如,可以为用户提供图片归纳能力(如,可以将手机相册中的同类照片进行归纳),或者,还可以用于区分原版图像与PS等方式处理过的图像,等等。

[0105] 下面对本申请实施例提供的具体实现方案进行详细介绍。

[0106] 实施例一

[0107] 首先,该实施例一提供了一种对象识别方法,参见图2,该方法可以包括:

[0108] S201:确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息。

[0109] 其中,待识别的目标对象组,可以根据实际的识别需求而确定的。例如,在商品对象信息系统的场景下,如前文所述,如果需要根据用户历史浏览商品,进行同款/相似款商品对象的推荐,则可以将历史浏览商品与待推荐商品池中的各个商品分别进行组合,得到多个商品对,每个商品对就可以成为一个商品对象组。或者,在根据页面中当前展示的某个商品对象A执行“找相似”时,可以将该商品对象A与商品库中的商品对象组成多个商品对象组(每个商品对象组中包括该商品对象A以及商品库中的另一个商品对象)。另外,在根据

用户上传的照片等进行正品判断等场景中,则可以将该照片中的拍摄对象与商品库中的商品对象组成多个商品对象组,分别进行识别,等等。

[0110] 其中,在确定出目标对象组之后,可以分别为各个目标对象采集出文本信息以及图像信息。其中,如果具体的目标对象都是系统中已发布的商品对象,则商品信息库中都可以保存为具体的商品对象关联的文本信息以及图像信息等,其中,文本信息可以通过商品对象的标题、商标等信息进行获得,图像信息则可以通过商品对象的主图等进行获得。而如果具体的目标对象包括用户通过上传图像信息、文本信息等方式指定的待识别对象,则可以直接从用户上传的信息中提取文本信息以及图像信息。其中,如果用户仅上传了图像信息,其中可能存在不包含文本信息的情况,因此,还可以提供报警功能,也即,在分别读入各待识别对象的文本、图像等内容过程中,如果发现某部分内容有缺失,则可以触发报警,提醒用户是否信息读入错误还是忽略继续读入,或者,还可以为用户提供输入文本信息的操作控件,以用于为具体上传的图像输入关联的文本内容,等等。

[0111] S202:将所述目标对象组中各对象的文本信息以及图像信息输入到目标模型中进行处理,得到对象识别结果。

[0112] 在确定出目标对象组,以及具体目标对象关联的文本内容以及图像内容之后,就可以利用预先训练完成的目标模型,对组内的各目标对象是否在目标属性上是否存在共性进行判断,例如,同组内的各商品对象是否属于同款或者相似款,某些照片是否属于同一场景,等等。

[0113] 其中,目标模型可以用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的各目标对象的图文交叉特征,判断各目标对象在目标属性上是否存在共性。在一种具体的实现方式下,具体的目标模型中可以包括特征生成模型以及判别模型,具体的特征生成模型用于:分别对各目标对象进行文本特征提取以及图像特征提取,并将文本特征与图像特征进行融合,输出各目标对象的图文交叉特征;所述判别模型用于根据各目标对象的图文交叉特征判断各目标对象在目标属性上是否存在共性。而关于具体的特征生成模型,具体可以是由文本特征提取模型,图像特征提取模型,以及多模态特征融合模型等组成的,文本特征提取模型主要用于从目标对象的文本内容中进行文本特征提取,图像特征提取模型用于从目标对象的图像内容中进行图像特征提取,分别得到文本特征向量,以及图像特征向量。多模态特征融合模型则用于对文本特征向量以及图像特征向量进行融合处理,得到图文交叉特征向量。

[0114] 其中,如前文所述,在本申请实施例中,多模态特征融合模型可是通过有监督学习的方式训练生成的。为了通过有监督学习的方式进行对多模态特征融合模型的训练,以提升特征融合结果的质量,可以将多模态特征融合模型的训练,转化为对目标对象的图像与文本是否匹配进行分类的模型的训练。也就是说,具体的多模态特征融合模型,是在执行对目标对象的文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,并将融合处理结果确定为所述图文交叉特征。

[0115] 这样,在对多模态特征模型进行训练时,可以预先建立起图文匹配分类模型,以便对图文匹配分类模型进行训练,完成对图文匹配分类模型的训练之后,再将该模型中的输出层去掉,就可以得到多模态特征模型。

[0116] 其中,具体对图文匹配分类模型进行训练时,可以获取到多条训练样本,以及对应

的标注信息。其中,每条训练样本中可以对应一个(图像,文本)对,标注信息就是具体的图像与文本是否匹配。例如,在商品对象场景下,具体实现时,可以预先根据商品对象信息系统中已发布的多个商品对象,将其主图与对应的标题等文本提取出来作为训练样本,由人工标注等方式对图像与文本是否匹配进行标注。当然,如果是系统中已发布的同一商品对象的图像以及文本,则两者通常是匹配的,但是,在具体训练时,通常还需要一些负样本,也即,图像与文本不匹配的样本。为此,具体实现时,如图3所示,还可以获取首先预标注的数据集D(其中包括多条图像以及对应的文本,并且其中大部分条码中图像与文本都是相匹配的情况),然后,通过将数据集D中的每个文本和图像进行若干次采样,得到数据集G。例如,原来的数据集D中包括(图像1,文本1),(图像2,文本2),(图像3,文本3),等等,通过若干次采样之后,使得数据集G中不仅可以包括上述数据条目,还可以包括(图像1,文本2),(图像1,文本3),(图像2,文本1),(图像2,文本3),(图像3,文本1),(图像3,文本2),等等。通过这种方式,不仅可以构造出更大数量的样本,并且还可以使得样本中不仅可以包括正样本,还可以包括负样本。例如,数据集G共n个样本,采样得到的图文匹配的样本标记为1,图文不匹配的样本标记为0,等等。

[0117] 在得到上述数据集G之后,可以对每个样本分别通过文本编码和Resnet等算法提取,得到文本特征序列和图像特征序列。例如,对输入的文本内容进行文本分词,去除停用词等操作后,采用预定义的分词等算法对文本进行编码,以此提取文本特征序列。另外,还可以对输入的图像切割成N小块(N可以自行设定,例如,可以设置为64),每块小图像采用预定义的Resnet等算法提取特征,得到图像特征序列。之后,可以利用Bert等算法,实现对文本特征序列和图像特征序列的融合,并且可以利用Relu(Rectified Linear Unit,线性整流函数)等作为激活函数(也就是具体执行融合计算的函数,主要用于对文本特征序列和图像特征序列中的数据进行计算得到映射结果,也即融合结果)。得到的融合结果可以用于对文本与图像是否匹配进行判断,并通过带入损失函数,并利用梯度下降等算法对模型中的参数W1进行求解。通过多轮迭代直到算法收敛后,可以将模型中的参数W1的取值确定下来,以此完成对模型的训练。

[0118] 完成对图文匹配分类模型的训练后,该模型的作用就是,输入一个目标对象的图像以及文本,则分类模型可以给出该目标对象的图像与文本是否匹配的结果。而本申请实施例中需要的是目标对象的图像特征与文本特征相融合之后的图文交叉特征,因此,还可以将该分类模型中的输出层去掉,该图文匹配分类模型就可以作为多模态特征融合模型来使用。也即,可以根据所述图文匹配模型中的特征融合模块,获得所述多模态特征融合模型。这样,在同样以一个目标对象的图像以及文本内容作为输入的情况下,其输出可以是图像特征与文本特征相融合之后的图文交叉特征。

[0119] 这样,具体在确定出目标对象组之后,可以将组内的各个目标对象分别输入到特征生成模型中,分别对各个目标对象的文本特征以及图像特征进行特征提取,并且可以得到各个目标对象的图文交叉特征。之后,可以用各目标对象的图文交叉特征,对各目标对象在目标属性上是否存在共性进行判断。

[0120] 其中,具体根据图文交叉特征对目标对象在目标属性上是否存在共性进行判断的判别模型可以有多种形式。例如,一种简单的方式下,具体的判别模型可以是对各目标对象的图文交叉特征向量进行距离计算(例如,向量间的欧式距离等),然后,根据具体计算出的

距离以及预先设置的阈值,判断出两个目标对象是否属于同款/相似款商品对象。

[0121] 为了进一步提升识别的准确度,在本申请的优选实施例中,关于判别模型,也可以是通过有监督学习的方式训练生成的。例如,可以预先构建起判别模型,该判别模型可以以多个目标对象的图文交叉特征为输入,输出多个目标对象是否在目标属性上是否存在共性(例如,各商品对象是否为同款或者相似款,等等)的判断结果。

[0122] 具体在对上述判别模型进行有监督学习的训练时,可以预先获取多个目标对象组(例如,商品对象A与商品对象B组成一个商品对象对,等等)作为训练样本,并且可以预先获取到商品对象组中各个目标对象的图文交叉特征(例如,预先利用之前训练好的特征生成模型来获得),并添加关于同一组内各目标对象是否在目标属性上存在共性的标注信息。然后,利用这种训练样本对判别模型进行训练,以使得向该判别模型输入一组目标对象的图文交叉特征后,可以输出关于这组目标对象是否在目标属性上存在共性的判别结果。

[0123] 另外,由于具体的判别模型是以目标对象的图文交叉特征作为输入进行判别,因此,实际上在判别之前,会涉及到对各个目标对象进行文本特征及图像特征的提取以及融合,以生成图文交叉特征的过程。具体实现时,关于特征生成模型以及上述判别模型可以单独进行有监督学习的训练,或者,在本申请的优选实施例中,为了提升训练效率,以及进一步提升最终模型识别的准确度,还可以直接将特征生成模型与判别模型组合在一起(例如,在特征生成模型之后拼接一个判别模型,等等),以此生成一整套的目标模型,该目标模型可以直接以多个目标对象的文本以及图像内容为输入,输出多个目标对象在目标属性上是否存在共性的判断结果。

[0124] 另外,在这种组合为一整套的目标模型的情况下,还可以对特征生成模型与判别模型一起进行训练。例如,具体在进行训练时,如图4所示,可以直接获取多个商品对象组(例如,商品对象1与商品对象2组成一个商品对象对,等等)作为训练样本,并添加关于同一组内各商品对象是否为同款或相似款的标注信息。之后,可以通过特征生成模型,分别对商品对象1以及商品对象2进行文本特征提取以及图像特征提取,再进行融合后生成图文交叉特征向量,分别为向量M和向量N。之后,该向量M和向量N可以输入到判别函数中进行判断,输出的判断结果带入损失函数,并利用梯度下降等算法对模型中的参数进行求解,等等。

[0125] 其中,为了进一步提升模型的性能,还可以首先通过有监督学习的方式单独对特征生成模型进行训练,并根据训练结果获得所述特征生成模型中的参数学习结果;例如,前述例子中的参数W1,通过单独对特征生成模型进行训练,可以将该W1的值确定下来。之后,再重新对特征生成模型与判别模型组成的一整套的目标模型进行有监督学习,并且,可以将之前单独训练特征生成模型时确定下来的参数值,作为所述目标模型中的参数初始值,然后再通过有监督学习的方式对所述目标模型进行训练。也就是说,在对上述一整套的目标模型进行训练时,不再使用随机的参数初始值,而是利用单独训练特征生成模型过程中确定下来的参数值作为初始值,这样可以进一步提升模型的性能,使得最终的识别准确度得到提升。

[0126] 总之,通过本申请实施例,可以通过将目标对象的多模态的特征(例如,包括文本信息与图像信息等)进行融合的方式,获得目标对象的图文交叉特征,然后可以基于这种图文交叉特征对目标对象在目标属性上是否存在共性进行更准确的识别。其中,在对多模态特征进行融合时,考虑到图像特征与文本特征属于不同域的情况,并不是直接将图像特征

与文本特征进行拼接,或者进行简单的数学运算,而是可以构建多模态特征融合模型,并通过有监督学习方式对该模型进行训练,从而可以通过该多模态特征融合模型,对文本特征以及图像特征进行多模态特征的融合。为了实现对多模态特征融合模型进行有监督学习方式的训练,还可以将其转化为对图文匹配分类模型的训练,也即,具体的多模态特征融合模型可以是在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,进而可以将融合处理结果确定为图文交叉特征。通过这种方式,可以生成有监督的多模态特征融合模型,以此生成更能表达目标对象特征的图文交叉特征,进而可以更准确地识别出目标对象在目标属性上是否存在共性(例如,各商品对象是否属于同款/相似款)。

[0127] 另外,在实际应用中,在一些场景下,还可以向用户提供关于各目标对象在目标属性上是否存在共性的识别结果,并且,可以向用户提供对识别结果进行反馈的操作选项。例如,在“找相似”场景中,在用户针对某商品对象A发起“找相似”请求后,通过模型计算可以返回识别结果,例如,可以包括商品对象B、C、D等等,还可以针对各个商品对象给出“反馈”操作选项,假设用户发现商品对象C与商品对象A并不是同款,则可以提交反馈信息,等等。在接收关于所述识别结果准确性的用户反馈信息后,还可以利用这种用户反馈信息对所述目标模型进行更新迭代训练。也就是说,用户反馈的信息也可以成为带标注的样本,可以参与到对模型的迭代训练中。

[0128] 这里需要说明的是,在具体实现时,在商品对象作为目标对象的场景中,还可以根据多个不同的商品对象类目,分别训练不同的目标模型,以用于分别使用具体类目的目标模型进行更准确的同款/相似款商品对象识别。这样,可以根据所述目标商品对象组中各商品对象所属的共同的类目,将各商品对象的文本信息以及图像信息输入到所述类目对应的目标模型中进行处理。例如,在具体需要针对某个商品对象A找同款或者相似款商品对象时,可以首先确定出该商品对象A所属的类目(商品对象信息库中会保存具体商品对象所属的类目信息,因此,可以通过查询商品对象信息库的方式获知商品对象的类目信息),并从商品库中首先根据类目信息进行商品对象的初筛,然后,利用该类目对应的目标模型,对该商品对象A与其同类目的其他商品对象是否为同款/相似款进行识别,等等。

[0129] 实施例二

[0130] 该实施例二是与实施例一相对应的,针对具体对多模态特征融合模型进行训练的过程中,提供了一种模型处理方法,参见图5,该方法可以包括:

[0131] S501:构建图文匹配模型,所述图文匹配模型用于对图像与文本是否用于描述统一对象进行识别;

[0132] S502:获取训练样本,所述训练样本中包括多个样本条目,所述样本条目中包括图像内容以及文本内容,以及关于图像内容与文本内容是否匹配的标注信息;

[0133] S503:通过所述训练样本对所述图文匹配模型进行训练;

[0134] S504:在训练完成后,根据所述图文匹配模型中的特征融合模块,生成所述多模态特征融合模型,以用于对输入的目标对象进行图像特征提取以及文本特征提取,并将图像文本与文本特征进行融合,生成图文交叉特征。

[0135] 实施例三

[0136] 该实施例三也是与实施例一相对应的,针对对目标模型(由特征生成模型以及判

别模型组合而成)的训练过程中,一种模型处理方法,参见图6,该方法可以包括:

[0137] S601:对所述特征生成模型进行训练,确定其中的参数值;所述特征生成模型用于将目标对象的图像特征以及文本特征进行融合,生成图文交叉特征;

[0138] S602:获取训练样本,所述训练样本中包括多个目标对象组,每个目标对象组中包括至少两个目标对象,以及关于所述至少两个目标对象是否在目标属性上是否存在共性的标注信息;

[0139] S603:利用所述训练样本对所述目标模型进行训练,并将对所述特征生成模型进行训练过程中确定的参数值,作为对所述目标模型进行训练时的参数的初始值。

[0140] 实施例四

[0141] 该实施例四针对其中一种具体的应用场景,提供了一种提供识别结果的方法,参见图7,该方法可以包括:

[0142] S701:接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

[0143] S702:将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

[0144] S703:将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,根据得到的图文交叉特征,确定目标对象之间相似度;其中,所述目标模型中包括多模态特征融合模型,所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征进行融合处理,以得到所述图文交叉特征;

[0145] S704:根据与所述待识别对象的相似度符合条件的目标商品对象对应的标签,提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0146] 关于上述实施例二至四中的未详述部分,可以参见实施例一以及本说明书中其他部分的记载,这里不再赘述。

[0147] 需要说明的是,本申请实施例中可能会涉及到对用户数据的使用,在实际应用中,可以在符合所在国的适用法律法规要求的情况下(例如,用户明确同意,对用户切实通知,等),在适用法律法规允许的范围内在本文描述的方案中使用用户特定的个人数据。

[0148] 与实施例一相对应,本申请实施例还提供了一种对象识别装置,参见图8,该装置可以包括:

[0149] 目标对象组确定单元801,用于确定待识别的目标对象组,所述目标对象组中包括至少两个待识别的目标对象,所述目标对象关联有文本信息以及图像信息;

[0150] 识别处理单元802,用于将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理,得到对象识别结果;其中,所述目标模型用于,分别将各目标对象的文本特征与图像特征进行融合处理,并根据得到的各目标对象的图文交叉特征,判断各目标对象在目标属性上是否存在共性;

[0151] 所述目标模型中包括多模态特征融合模型,所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中,对所述文本特征与图像特征

进行融合处理,以得到所述图文交叉特征。

[0152] 具体实现时,所述目标模型包括特征生成模型以及判别模型,所述特征生成模型中包括特征提取模型,以及多模态特征融合模型;

[0153] 所述特征提取模型用于:分别对各目标对象进行文本特征提取以及图像特征提取;

[0154] 所述多模态特征融合模型用于:将文本特征与图像特征进行融合,输出各目标对象的图文交叉特征;

[0155] 所述判别模型用于根据各目标对象的图文交叉特征判断各目标对象在目标属性上是否存在共性。

[0156] 其中,所述目标模型是通过将特征生成模型以及判别模型进行组合而成的成套模型。

[0157] 具体实现时,在对所述目标模型进行训练之前,还可以对所述特征生成模型进行训练,并根据训练结果获得所述特征生成模型中的参数学习结果;然后,将所述参数学习结果作为对所述目标模型进行训练时的参数初始值。

[0158] 具体实现时,该装置还可以包括:

[0159] 结果反馈单元,用于提供关于目标对象在目标属性上是否存在共性的识别结果,并接收关于所述识别结果准确性的用户反馈信息,根据所述用户反馈信息对所述目标模型进行迭代训练。

[0160] 其中,所述目标对象包括商品对象;

[0161] 所述识别处理单元具体用于:确定同一商品对象组中的各商品对象是否属于同款或相似款商品对象。

[0162] 其中,所述商品对象组中包括的至少两个商品对象对应相同的类目;

[0163] 所述目标模型包括对应不同类目的多个不同目标模型;

[0164] 所述识别处理单元具体可以用于:

[0165] 根据所述目标商品对象组中各商品对象所属的目标类目,将所述各商品对象的文本信息以及图像信息输入到所述目标类目对应的目标模型中进行处理。

[0166] 其中,目标对象组确定单元具体可以用于:

[0167] 在根据用户历史浏览过的商品对象进行商品对象推荐时,将该历史浏览过的商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

[0168] 或者,目标对象组确定单元具体可以用于:

[0169] 所述确定待识别的目标对象组,包括:

[0170] 在接收到用户提交的查找目标商品对象的同款/相似款商品对象的请求后,将该目标商品对象与商品对象库中的其他数据对象组成多个目标商品对象组。

[0171] 或者,目标对象组确定单元具体可以用于:

[0172] 接收用户提交的待识别对象的文本信息以及图像信息,所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的;

[0173] 将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组;所述商品对象库中的商品对象还关联有标签,所述标签用于表征对应的商品对象是否为正品或与正品的接近程度;

[0174] 所述装置还可以包括：

[0175] 正品识别结果提供单元，用于根据与所述待识别对象相似度符合条件的目标商品对象对应的标签，提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0176] 与实施例二相对应，本申请实施例还提供了一种模型处理装置，参见图9，该装置可以包括：

[0177] 模型构建单元901，用于构建图文匹配模型，所述图文匹配模型用于对图像与文本是否用于描述统一对象进行识别；

[0178] 训练样本获取单元902，用于获取训练样本，所述训练样本中包括多个样本条目，所述样本条目中包括图像内容以及文本内容，以及关于图像内容与文本内容是否匹配的标注信息；

[0179] 训练单元903，用于通过所述训练样本对所述图文匹配模型进行训练；

[0180] 多模态特征融合模型生成单元904，用于在训练完成后，根据所述图文匹配模型中的特征融合模块，生成所述多模态特征融合模型，以用于对输入的目标对象进行图像特征提取以及文本特征提取，并将图像文本与文本特征进行融合，生成图文交叉特征。

[0181] 与实施例三相对应，本申请实施例还提供了一种模型处理装置，所述模型是由特征生成模型与判别模型进行组合而成的，参见图10，所述装置包括：

[0182] 第一训练单元1001，用于对所述特征生成模型进行训练，确定其中的参数值；所述特征生成模型用于将目标对象的图像特征与文本特征进行融合，生成图文交叉特征；

[0183] 训练样本获取单元1002，用于获取训练样本，所述训练样本中包括多个目标对象组，每个目标对象组中包括至少两个目标对象，以及关于所述至少两个目标对象是否在目标属性上是否存在共性的标注信息，其中，每个商品对象关联的文本信息以及图像信息；

[0184] 第二训练单元1003，用于利用所述训练样本对所述目标模型进行训练，其中，将所述对所述特征生成模型进行训练过程中确定的参数值，作为对所述目标模型进行训练时的参数的初始值。

[0185] 与实施例四相对应，本申请实施例还提供了一种提供识别结果的装置，参见图11，该装置可以包括：

[0186] 信息接收单元1101，用于接收用户提交的待识别对象的文本信息以及图像信息，所述图像信息是通过对待识别对象对应的实物进行图像采集后生成的；

[0187] 目标对象组生成单元1102，用于将所述待识别对象与商品对象库中的多个商品对象分别组成目标对象组；所述商品对象库中的商品对象还关联有标签，所述标签用于表征对应的商品对象是否为正品或与正品的接近程度；

[0188] 相似度识别单元1103，用于将所述目标对象组中各目标对象的文本信息以及图像信息输入到目标模型中进行处理，所述目标模型用于，分别将各目标对象的文本特征与图像特征进行融合处理，根据得到的图文交叉特征，确定目标对象之间相似度；其中，所述目标模型中包括多模态特征融合模型，所述多模态特征融合模型用于在执行对文本特征以及图像特征是否匹配进行判断的任务的过程中，对所述文本特征与图像特征进行融合处理，以得到所述图文交叉特征；

[0189] 识别结果提供单元1104，用于根据与所述待识别对象的相似度符合条件的目标商

品对象对应的标签,提供关于所述待识别对象是否为正品或与正品的接近程度的识别结果信息。

[0190] 另外,本申请实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现前述方法实施例中任一项所述的方法的步骤。

[0191] 以及一种电子设备,包括:

[0192] 一个或多个处理器;以及

[0193] 与上述一个或多个处理器关联的存储器,所述存储器用于存储程序指令,所述程序指令在被上述一个或多个处理器读取执行时,执行前述方法实施例中任一项所述的方法的步骤。

[0194] 其中,图12示例性的展示出了电子设备的架构,具体可以包括处理器1210,视频显示适配器1211,磁盘驱动器1212,输入/输出接口1213,网络接口1214,以及存储器1220。上述处理器1210、视频显示适配器1211、磁盘驱动器1212、输入/输出接口1213、网络接口1214,与存储器1220之间可以通过通信总线1230进行通信连接。

[0195] 其中,处理器1210可以采用通用的CPU(Central Processing Unit,中央处理器)、微处理器、应用专用集成电路(Application Specific Integrated Circuit,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,以实现本申请所提供的技术方案。

[0196] 存储器1220可以采用ROM(Read Only Memory,只读存储器)、RAM(Random Access Memory,随机存取存储器)、静态存储设备,动态存储设备等形式实现。存储器1220可以存储用于控制电子设备1200运行的操作系统1221,用于控制电子设备1200的低级别操作的基本输入输出系统(BIOS)。另外,还可以存储网页浏览器1223,数据存储管理系统1224,以及对象识别处理系统1225等等。上述对象识别处理系统1225就可以是本申请实施例中具体实现前述各步骤操作的应用程序。总之,在通过软件或者固件来实现本申请所提供的技术方案时,相关的程序代码保存在存储器1220中,并由处理器1210来调用执行。

[0197] 输入/输出接口1213用于连接输入/输出模块,以实现信息输入及输出。输入输出/模块可以作为组件配置在设备中(图中未示出),也可以外接于设备以提供相应功能。其中输入设备可以包括键盘、鼠标、触摸屏、麦克风、各类传感器等,输出设备可以包括显示器、扬声器、振动器、指示灯等。

[0198] 网络接口1214用于连接通信模块(图中未示出),以实现本设备与其他设备的通信交互。其中通信模块可以通过有线方式(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。

[0199] 总线1230包括一通路,在设备的各个组件(例如处理器1210、视频显示适配器1211、磁盘驱动器1212、输入/输出接口1213、网络接口1214,与存储器1220)之间传输信息。

[0200] 需要说明的是,尽管上述设备仅示出了处理器1210、视频显示适配器1211、磁盘驱动器1212、输入/输出接口1213、网络接口1214,存储器1220,总线1230等,但是在具体实施过程中,该设备还可以包括实现正常运行所必需的其他组件。此外,本领域的技术人员可以理解的是,上述设备中也可以仅包含实现本申请方案所必需的组件,而不必包含图中所示的全部组件。

[0201] 通过以上的实施方式的描述可知,本领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现。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可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如ROM/RAM、磁碟、光盘等,包括若干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或者实施例的某些部分所述的方法。

[0202] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统或系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统及系统实施例仅仅是示意性的,其中所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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0203] 以上对本申请所提供的对象识别方法及电子设备,进行了详细介绍,本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的一般技术人员,依据本申请的思想,在具体实施方式及应用范围上均会有改变之处。综上所述,本说明书内容不应理解为对本申请的限制。

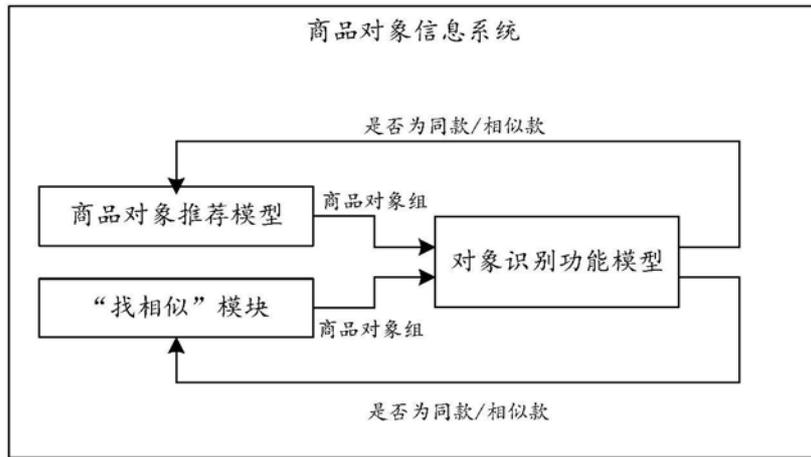


图1

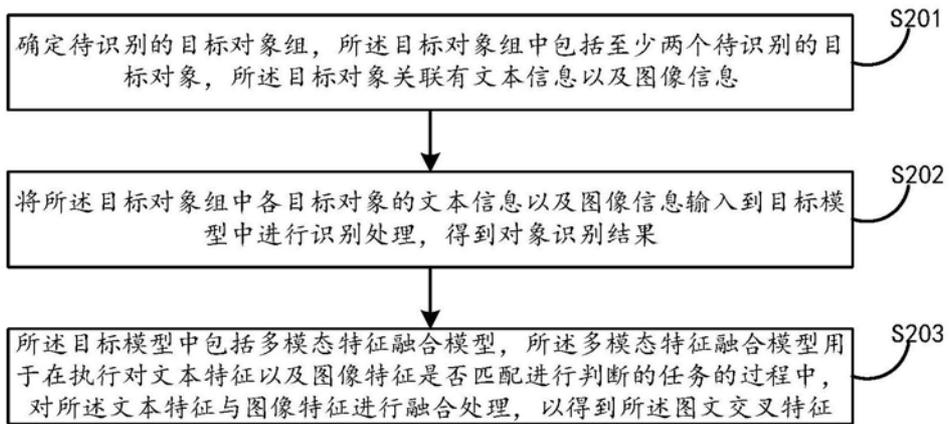


图2

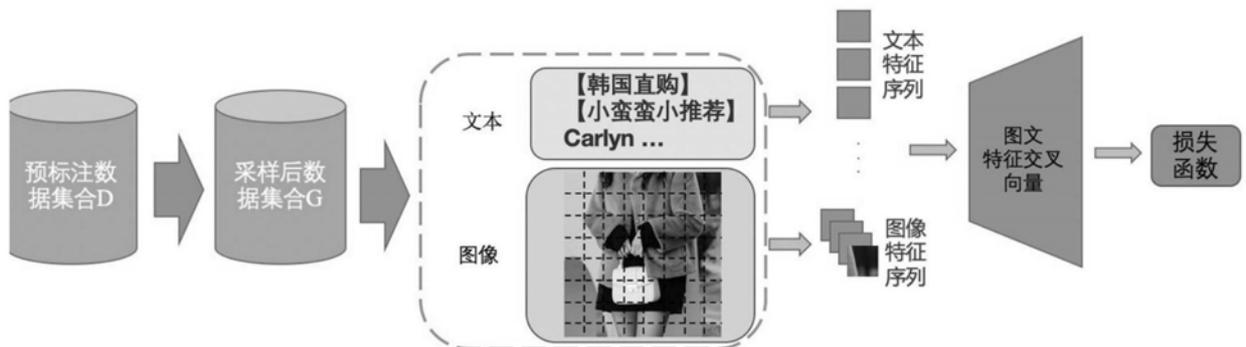


图3

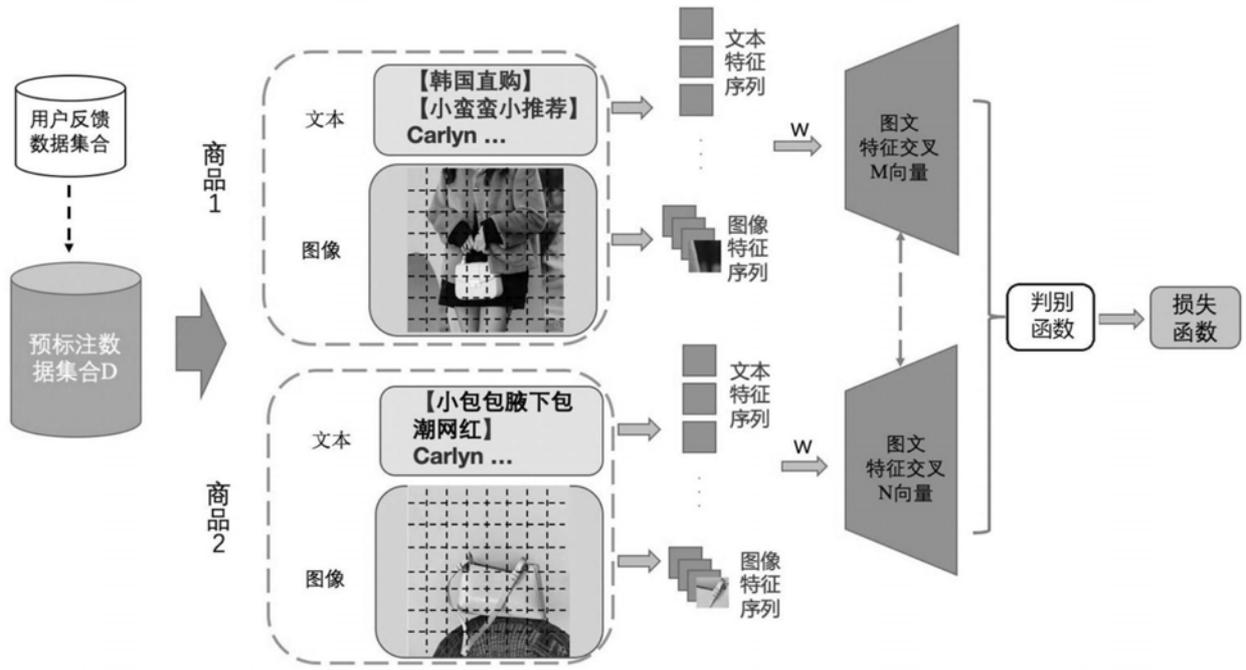


图4

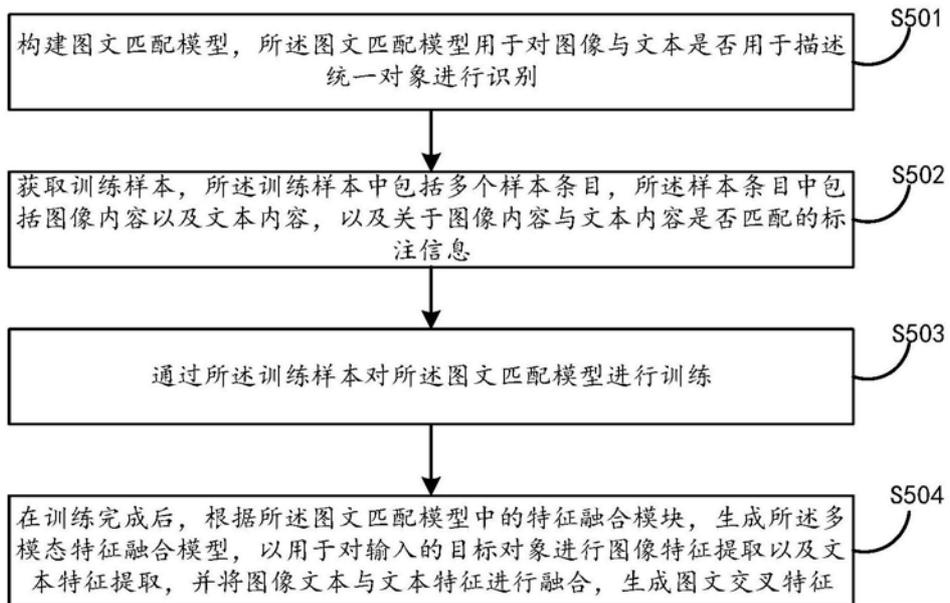


图5

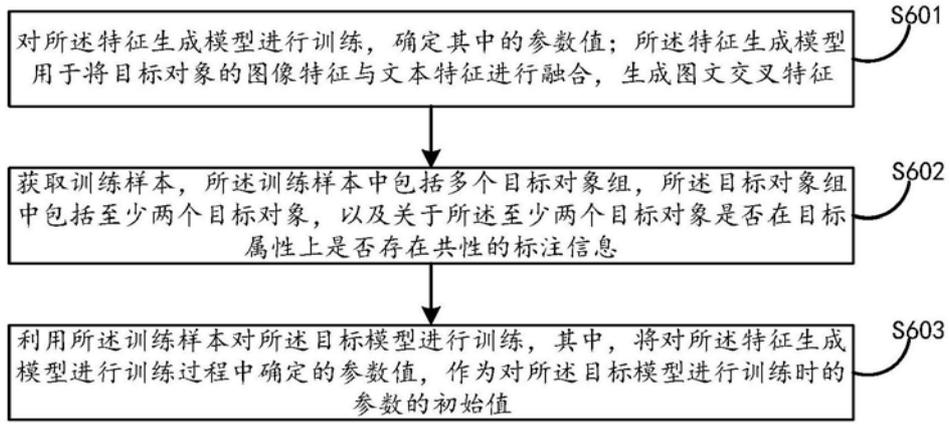


图6

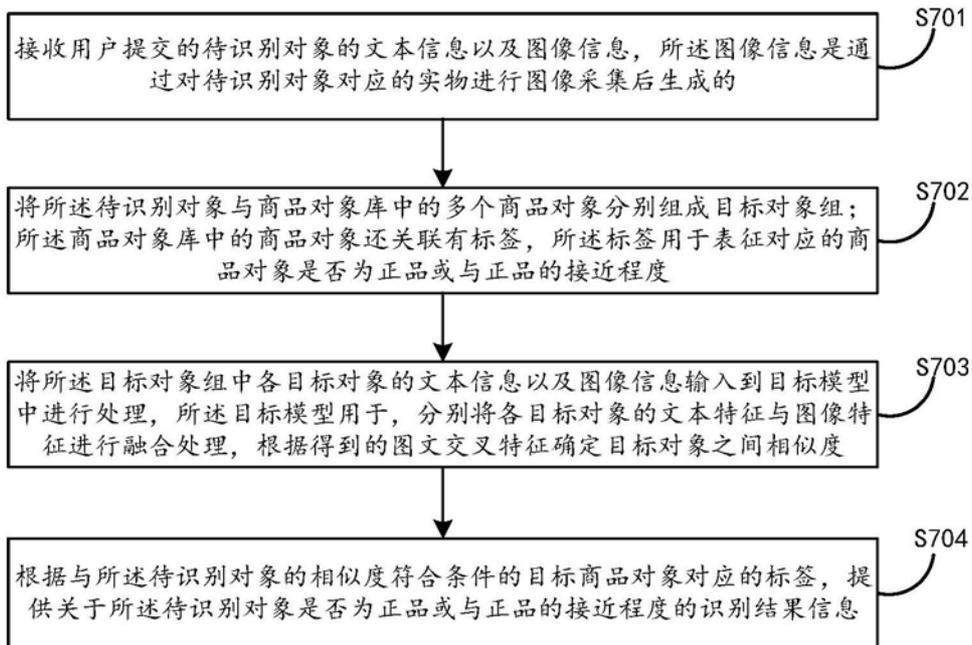


图7



图8

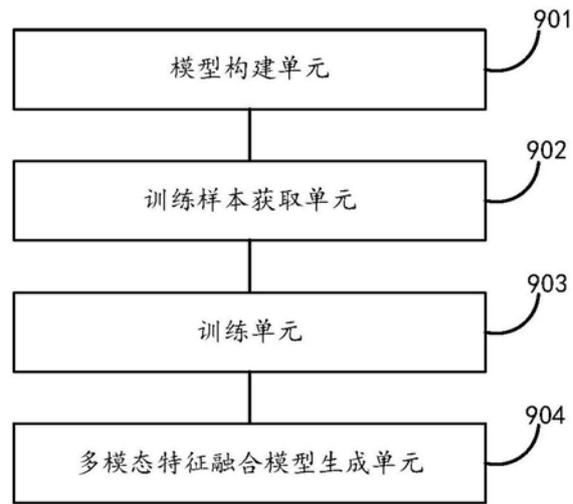


图9

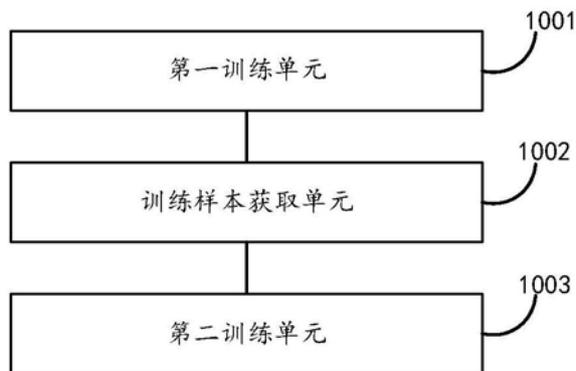


图10

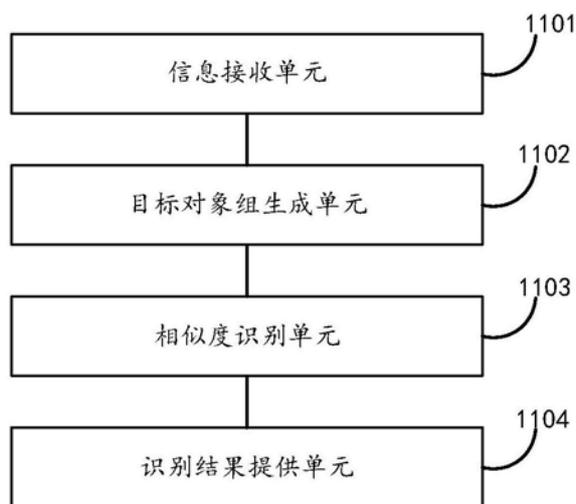


图11

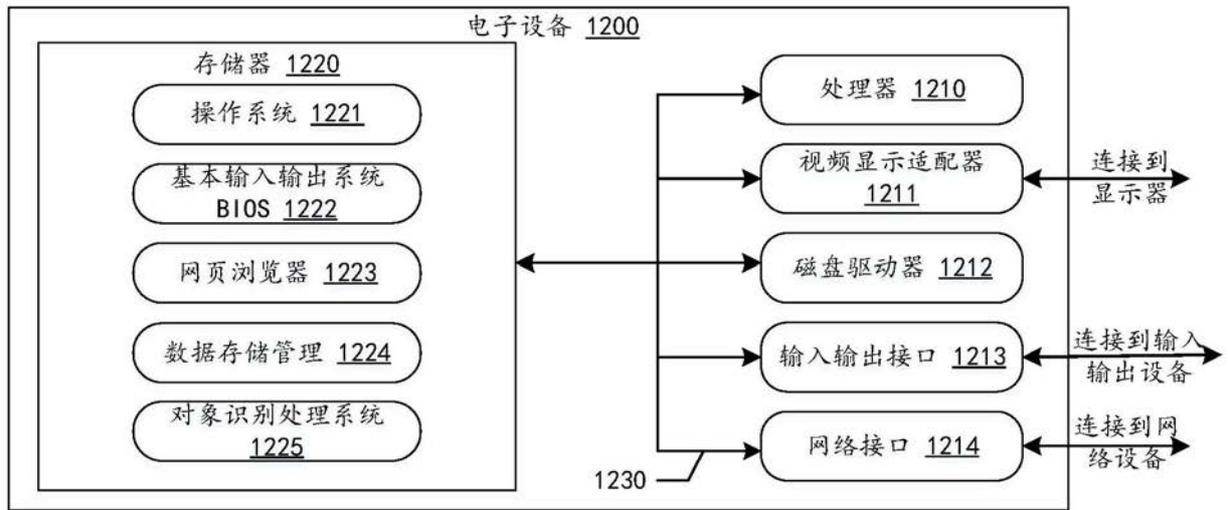


图12