



(12) 发明专利申请

(10) 申请公布号 CN 113762319 A

(43) 申请公布日 2021.12.07

(21) 申请号 202110239408.4

(22) 申请日 2021.03.04

(71) 申请人 北京沃东天骏信息技术有限公司
地址 100176 北京市大兴区经济技术开发区
科创十一街18号院2号楼4层A402室
申请人 北京京东世纪贸易有限公司

(72) 发明人 袁韶祖 陈蒙 刘瑞雪

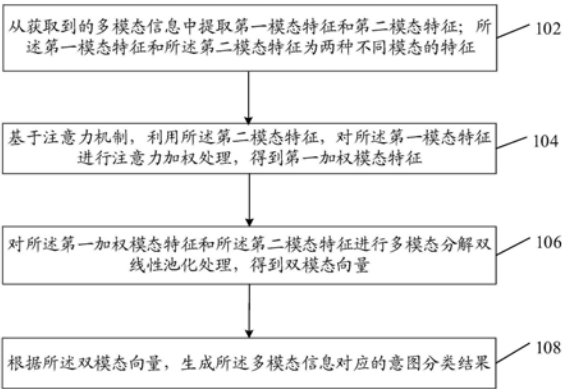
(74) 专利代理机构 北京派特恩知识产权代理有限公司 11270
代理人 李昂 张颖玲

(51) Int.Cl.
G06K 9/62 (2006.01)
G06F 40/30 (2020.01)
G06N 3/04 (2006.01)

权利要求书2页 说明书15页 附图4页

(54) 发明名称
一种多模态信息处理方法、装置、电子设备和存储介质

(57) 摘要
本申请实施例公开了一种多模态信息处理方法、装置、电子设备和存储介质,所述方法包括:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;根据所述双模态向量,生成所述多模态信息对应的意图分类结果。



1. 一种多模态信息处理方法,其特征在于,所述方法包括:

从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;

对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;

根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

2. 根据权利要求1所述的方法,其特征在于,所述第一模态特征和所述第二模态特征均为至少一个时,所述基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征,包括:

将至少一个第一模态特征和至少一个第二模态特征映射到第一空间维度;

针对每一所述第二模态特征,执行如下步骤:

在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征和所述第二模态特征之间的相关度;

根据每一所述相关度,确定对应第一模态特征的注意力权重;

利用每一所述注意力权重,对各第一模态特征进行加权求和,得到所述第二模态特征对应的第一加权模态特征。

3. 根据权利要求2所述的方法,其特征在于,所述在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征和所述第二模态特征之间的相关度,包括:

确定每一所述第一模态特征对应的第一特征向量和所述第二模态特征对应的第二特征向量;

确定每一所述第一特征向量和所述第二特征向量之间的点积;

根据每一所述点积,确定对应第一模态特征和所述第二模态特征之间的相关度。

4. 根据权利要求1所述的方法,其特征在于,所述第一加权模态特征和所述第二模态特征均为至少一个时,所述对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量,包括:

将每一所述第一加权模态特征和对应的所述第二模态特征映射到第二空间维度;

在所述第二空间维度下,确定所述第一加权模态特征对应的第三特征向量和所述第二模态特征对应的第四特征向量;

确定所述第三特征向量和所述第四特征向量之间的外积;

对所述外积进行求和池化,得到所述第一加权模态特征对应的双模态向量。

5. 根据权利要求1至4中任一项所述的方法,其特征在于,所述第一模态特征包括图像特征,所述第二模态信息包括文本特征,所述从获取到的多模态信息中提取第一模态特征和第二模态特征,包括:

利用卷积神经网络,从获取到的多模态信息中提取图像特征;

利用循环神经网络,从所述多模态信息中提取文本特征。

6. 根据权利要求1至4中任一项所述的方法,其特征在于,所述方法还包括:

确定与所述意图分类结果匹配的知识点;

根据所述知识点,生成多模态交互文本。

7. 根据权利要求6所述的方法,其特征在于,所述确定与所述意图分类结果匹配的知识点,包括:

确定与所述意图分类结果匹配的至少一个待选知识点;

显示所述至少一个待选知识点;

根据接收到的指令,从所述至少一个待选知识点中确定出目标知识点;

将所述目标知识点确定为与所述意图分类结果匹配的知识点。

8. 根据权利要求7所述的方法,其特征在于,所述显示所述待选知识点,包括:

确定每一所述待选知识点在历史时刻被确定为目标知识点的频率;

根据所述频率递减的顺序,对所述至少一个待选知识点进行排序;

按照排列顺序,显示所述至少一个待选知识点。

9. 根据权利要求6所述的方法,其特征在于,所述确定与所述意图分类结果匹配的知识点,包括:

确定与所述意图分类结果匹配的至少一个待选知识点;

确定每一所述待选知识点在历史时刻被确定为目标知识点的频率;

将频率最高的待选知识点,确定为与所述意图分类结果匹配的知识点。

10. 一种多模态信息处理装置,其特征在于,所述装置包括:

提取模块,用于从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

加权模块,用于基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;

池化模块,用于对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;

第一生成模块,用于根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

11. 一种电子设备,包括存储器和处理器,所述存储器存储有可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现权利要求1至9任一项所述多模态信息处理方法中的步骤。

12. 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,该计算机程序被处理器执行时实现权利要求1至9任一项所述多模态信息处理方法中的步骤。

一种多模态信息处理方法、装置、电子设备和存储介质

技术领域

[0001] 本申请涉及计算机技术,涉及但不限于一种多模态信息处理方法、装置、电子设备和存储介质。

背景技术

[0002] 相关技术中,多模态信息处理的解决方式主要有:基于单模态语言的分类或匹配模型的处理方式,如TextCNN(Text Convolutional Neural Networks,利用卷积神经网络进行文本分类)等;基于单模态图像分类模型的处理方式,如ResNet(Residual Network,残差网络)等。

[0003] 多模态信息处理中如果采用单模态(即只使用图像或者文本)方法进行解决,便浪费了其他模态的信息,得到的意图分类结果的准确率不高。

发明内容

[0004] 有鉴于此,本申请实施例提供一种多模态信息处理方法、装置、电子设备和存储介质。

[0005] 第一方面,本申请实施例提供一种多模态信息处理方法,所述方法包括:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0006] 第二方面,本申请实施例提供一种多模态信息处理装置,包括:提取模块,用于从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;加权模块,用于基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;池化模块,用于对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;第一生成模块,用于根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0007] 第三方面,本申请实施例提供一种电子设备,包括存储器和处理器,所述存储器存储有可在处理器上运行的计算机程序,所述处理器执行所述程序时实现本申请实施例任一所述多模态信息处理方法中的步骤。

[0008] 第四方面,本申请实施例提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现本申请实施例任一所述多模态信息处理方法中的步骤。

[0009] 本申请实施例中,通过基于注意力机制,利用第二模态特征,对第一模态特征进行注意力加权处理,得到第一加权模态特征,再对第一加权模态特征和第二模态特征进行

多模态分解双线性池化处理,得到双模态向量,进而生成意图分类结果,从而能够使得第一模态特征和第二模态特征进行充分的交互融合,获得更准确的意图分类结果。

附图说明

- [0010] 图1为本申请实施例一种多模态信息处理方法的流程示意图;
- [0011] 图2为本申请实施例一种基于自上而下注意力机制的神经网络模型的结构示意图;
- [0012] 图3为本申请实施例一种MFB的结构示意图;
- [0013] 图4为本申请实施例一种多模态信息的示意图;
- [0014] 图5为本申请实施例一种意图分类结果和知识点的匹配关系示意图;
- [0015] 图6为本申请实施例一种意图卡片的示意图;
- [0016] 图7为本申请实施例一种意图分类结果的生成方法示意图;
- [0017] 图8为本申请实施例一种多模态信息处理装置的组成结构示意图;
- [0018] 图9为本申请实施例电子设备的一种硬件实体示意图。

具体实施方式

[0019] 下面结合附图和实施例对本申请的技术方案进一步详细阐述。

[0020] 图1为本申请实施例提供的多模态信息处理方法的实现流程示意图,应用于电子设备,如图1所示,该方法包括:

[0021] 步骤102:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

[0022] 其中,每一种信息的来源或者形式,都可以称为一种模态;多模态信息可以由至少两种不同来源或者不同形式的信息组成的信息;所述多模态信息可以包含图像信息、音频信息和文本信息等信息中的至少两种;对应地,所述第一模态特征和所述第二模态特征可以是图像特征、音频特征或文本特征等,且所述第一模态特征和所述第二模态特征是不同的模态的特征。

[0023] 步骤104:基于注意力机制,利用所述第二模态特征,对所述第一模态特征进行注意力加权处理,得到第一加权模态特征;

[0024] 其中,在认知科学中,由于信息处理的瓶颈,会选择性地关注所有信息的一部分,同时忽略其他可见的信息。上述机制通常被称为注意力机制。所述注意力机制可以是自上而下的有意识的注意力(Top Down Attention),称为聚焦式注意力;聚焦式注意力是指有预定目的、依赖任务的、主动有意识地聚焦于某一对象的注意力。所述注意力机制还可以是自下而上(Bottom Up Attention)的无意识的注意力,称为基于显著性的注意力。基于显著性的注意力是由外界刺激驱动的注意,不需要主动干预,也和任务无关。

[0025] 所述第一模态特征的数量可以为至少一个;注意力加权处理可以是利用所述第二模态特征,为每一所述第一模态特征赋予对应的注意力权重,并根据每一所述第一模态特征的注意力权重,对各第一模态特征进行加权求和处理,得到第一加权模态特征。

[0026] 步骤106:对所述第一加权模态特征和所述第二模态特征进行多模态分解双线性池化处理,得到双模态向量;

[0027] 其中,多模态双线性池化(Multi-modal Factorized Bilinear Pooling,MFB) 处理可以用于对两个模态的特征进行特征融合;所述双模态向量为第一加权模态特征和第二模态特征融合后的向量。

[0028] 步骤108:根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0029] 其中,可以将双模态向量输入到循环神经网络中,利用所述循环神经网络对双模态向量进行训练,生成上下文向量;并可以利用上下文向量,生成意图分类结果;假设多模态信息为用户发送的文本信息和图像信息,在所述文本信息和图像信息的内容为咨询价保类问题的情况下,意图分类结果可以是“订单及物流截图”;在所述文本信息的内容为“帮我定个炸鸡送到XX酒店”、图像信息的内容为炸鸡的图片的情况下,意图分类结果可以是“订餐”。

[0030] 本申请实施例中,不仅基于注意力机制和第二模态特征,对第一模态特征进行注意力加权处理,得到第一加权模态特征,还对第一加权模态特征和第二模态特征进行多模态分解双线性池化处理,得到双模态向量,进而生成意图分类结果,从而能够使得第一模态特征和第二模态特征的每个点都进行了交互,融合效果更好。

[0031] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0032] 步骤S202:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

[0033] 其中,所述第一模态特征和所述第二模态特征均为至少一个;

[0034] 步骤S204:将至少一个第一模态特征和至少一个第二模态特征映射到第一空间维度;

[0035] 其中,假设第一模态特征为M维的特征,第二模态特征为N维的特征,所述第一空间维度为P维的空间,则可以将M维的第一模态特征和N维的第二模态特征映射到P维的空间;即所述第一模态特征由M维特征变为P维特征,所述第二模态特征由N维特征变为P维特征。

[0036] 针对每一所述第二模态特征,执行如下步骤S206至步骤S210:

[0037] 步骤S206:在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征和所述第二模态特征之间的相关度;

[0038] 步骤S208:根据每一所述相关度,确定对应第一模态特征的注意力权重;

[0039] 其中,所述相关度和所述注意力权重可以成正比,即第一模态特征和第二模态特征的相关度越高,第一模态特征的注意力权重越大。

[0040] 步骤S210:利用每一所述注意力权重,对各第一模态特征进行加权求和,得到所述第二模态特征对应的第一加权模态特征。

[0041] 可以采用步骤S206至步骤S210记载的方法,确定每一第二模态特征对应的第一加权模态特征。

[0042] 步骤S212:对每一所述第一加权模态特征和对应第二模态特征进行多模态分解双线性池化处理,得到双模态向量;

[0043] 步骤S214:根据每一所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0044] 其中,可以对各双模态向量进行处理,得到处理完成的双模态向量,并根据所述

处理完成的双模态向量,生成所述多模态信息对应的意图分类结果。

[0045] 本申请实施例中,通过第一模态特征和第二模态特征之间的相关度,确定第一模态特征的注意力权重,从而能够使得注意力权重的确定更准确,使得第一模态特征和第二模态得到一定程度的融合。

[0046] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0047] 步骤S302:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

[0048] 其中,所述第一模态特征和所述第二模态特征均为至少一个;

[0049] 步骤S304:将至少一个第一模态特征和至少一个第二模态特征映射到第一空间维度;

[0050] 其中,假设第一模态特征为图像特征,第二模态特征为文本特征,所述注意力机制为自上而下注意力(Top Down Attention)机制,图2为本申请实施例一种基于自上而下注意力机制的神经网络模型的结构示意图,参见图2,W可以表示为全连接层,softmax表示softmax层,可以先将图像特征201和文本特征202映射到512维的空间维度,再将图像特征和文本特征映射到2048维的空间维度。

[0051] 针对每一所述第二模态特征,执行如下步骤S306至步骤S314:

[0052] 步骤S306:在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征对应的第一特征向量和所述第二模态特征对应的第二特征向量;

[0053] 步骤S308:确定每一所述第一特征向量和所述第二特征向量之间的点积;

[0054] 步骤S310:根据每一所述点积,确定对应第一模态特征和所述第二模态特征之间的相关度。

[0055] 其中,点积和相关度可以成正比,即点积越大,可以认为第一特征向量和第二特征向量的相关度越大,对应地,第一模态特征和第二模态特征越相似;参见图2,可以使用softmax层的softmax函数对点积进行归一化处理,得到第一模态特征和第二模态特征之间的相关度。

[0056] 步骤S312:根据每一所述相关度,确定对应第一模态特征的注意力权重;

[0057] 其中,相关度和注意力权重可以呈正比,即相关度越大,可以认为第一模态特征的注意力权重越大。

[0058] 步骤S314:利用每一所述注意力权重,对各第一模态特征进行加权求和,得到所述第二模态特征对应的第一加权模态特征。

[0059] 其中,参见图2, Σ 可以表示求和操作,k可以表示第一模态特征的数量,则可以利用确定出的第一模态特征的注意力权重,对k个第一模态特征进行加权求和操作,得到第二模态特征对应的第一加权模态特征;同理,可确定出至少一个第二模态特征中每一第二模态特征对应的第一加权模态特征。

[0060] 步骤S316:对每一所述第一加权模态特征和对应第二模态特征进行多模态分解双线性池化处理,得到双模态向量;

[0061] 步骤S318:根据每一所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0062] 本申请实施例中,通过根据第一特征向量和第二特征向量之间的点积,确定第一

模态特征和第二模态特征之间的相关度,从而能够使得相关度的确定更加准确。

[0063] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0064] 步骤S402:从获取到的多模态信息中提取第一模态特征和第二模态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

[0065] 其中,所述第一模态特征和所述第二模态特征均为至少一个;

[0066] 步骤S404:将至少一个第一模态特征和至少一个第二模态特征映射到第一空间维度;

[0067] 针对每一所述第二模态特征,执行如下步骤S406至步骤S422:

[0068] 步骤S406:在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征对应的第一特征向量和所述第二模态特征对应的第二特征向量;

[0069] 步骤S408:确定每一所述第一特征向量和所述第二特征向量之间的点积;

[0070] 步骤S410:根据每一所述点积,确定对应第一模态特征和所述第二模态特征之间的相关度。

[0071] 步骤S412:根据每一所述相关度,确定对应第一模态特征的注意力权重;

[0072] 步骤S414:利用每一所述注意力权重,对各第一模态特征进行加权求和,得到所述第二模态特征对应的第一加权模态特征。

[0073] 步骤S416:将所述第二模态特征和对应的第一加权模态特征映射到第二空间维度;

[0074] 其中,假设第一加权模态特征为m维的特征,所述第二模态特征为n维的特征,则所述第二空间维度的可以为o维的空间维度;图3为本申请实施例一种MFB的结构示意图,参见图3,假设第一加权模态特征为加权图像特征301,加权图像特征301可以表示为 $x \in R^m$;第二模态特征为文本特征302,文本特征302可以表示为 $y \in R^n$,多模态分解式双线性池模型的输出可以表示为 $Z_i \in R$,则 Z_i 的维度可以为o;将所述第一加权模态特征映射到第二空间维度后可以得到第二空间维度下的第一加权模态特征303,将所述第二模态特征映射到第二空间维度后可以得到第二空间维度下的第二模态特征304。

[0075] 步骤S418:在所述第二空间维度下,确定所述第一加权模态特征对应的第三特征向量和第二模态特征对应的第四特征向量;

[0076] 步骤S420:确定所述第三特征向量和第四特征向量之间的外积;

[0077] 其中,多模态分解式双线性池模型可以定义为如下公式(1)所示:

$$[0078] \quad Z_i = x^T w_i y \quad \text{公式(1);}$$

[0079] 其中, $w_i \in R^{m \times n}$ 是一个映射矩阵, $Z_i \in R$ 是多模态分解式双线性池模型的输出,为了得到维度为o的输出,所有映射矩阵 $w = [w_1, \dots, w_o] \in R^{m \times n \times o}$,多模态分解式双线性池模型可以转换为如下公式(2)所示:

$$[0080] \quad Z_i = x^T u_i v_i^T y = \sum_{d=1}^k x^T u_d v_d^T y = I^T (x u_i^T o v_i^T y) \quad \text{公式(2);}$$

[0081] 其中,k是因子或因式化矩阵的潜在维度,I表示单位矩阵, $u_i = [u_1, \dots, u_k] \in R^{m \times k}$, $v_i = [v_1, \dots, v_k] \in R^{n \times k}$, $x u_i^T o v_i^T y$ 可以是第三特征向量和第四特征向量之间的外积305,又

可以称为Hadamard (哈达玛积) 积或第三特征向量和 第四特征向量的元素方向乘法,要学习的注意力权重为 $u=[u_1, \dots, u_k] \in R^{m \times k \times x_o}$, $v=[v_1, \dots, v_k] \in R^{n \times k \times x_o}$ 。

[0082] 步骤S422:对所述外积进行求和池化,得到所述第一加权模态特征对应的 双模态向量;

[0083] 其中,可以使用SumPooling函数对所述外积305进行求和池化;可以将U 和V改写成 $\tilde{u} \in R^{m \times k \times o}$, $\tilde{v} \in R^{n \times k \times o}$,则双模态向量306可以使用如下公式 (3) 表示:

[0084] $Z = \text{SumPooling}(\tilde{u}^\top x o \tilde{v}^\top y, k)$ 公式 (3);

[0085] 其中,函数 $\text{SumPooling}(\tilde{u}^\top x o \tilde{v}^\top y, k)$ 表示使用大小为k的一维非重叠窗口 对 $\tilde{u}^\top x o \tilde{v}^\top y$ 执行SumPooling。

[0086] 假设 $x_1 \in R^{k_o}$, $y_1 \in R^{k_o}$, $x_1 = \tilde{u}^\top x$, $y_1 = \tilde{v}^\top y$,则公式 (3) 可以转换为:

[0087] $Z = \text{SumPooling}(x_1 o y_1, k)$ 公式 (4);

[0088] 可以采用步骤S406至步骤S422记载的方法,确定每一第二模态特征对应 的第一加权模态特征,进而确定对应的双模态向量。

[0089] 步骤S424:根据每一所述双模态向量,生成所述多模态信息对应的意图分 类结果。

[0090] 本申请实施例中,在利用第二模态特征,对第一模态特征进行注意力加权 处理,得到第一加权模态特征的基础上,通过对第一加权模态特征和第二模态 特征进行多模态分解双线性池化处理,从而能够使得第一模态特征和第二模态 特征的每个点都进行了交互,融合效果更好。

[0091] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0092] 步骤S502:从获取到的多模态信息中提取图像特征和文本特征;

[0093] 步骤S504:基于注意力机制,利用所述文本特征,对所述图像特征进行注 意力加权处理,得到加权图像特征;

[0094] 步骤S506:对所述加权图像特征和所述文本特征进行多模态分解双线性池 化处理,得到双模态向量;

[0095] 步骤S508:根据所述双模态向量,生成所述多模态信息对应的意图分类结 果。

[0096] 本申请实施例中,不仅通过基于注意力机制和文本特征,对图像特征进行 注意力加权处理,得到加权图像特征,还对加权图像模态特征和文本特征进行 多模态分解双线性池化处理,得到双模态向量,进而生成意图分类结果,从而 能够使得图像特征和文本特征的每个点都进行了交互,融合效果更好。

[0097] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0098] 步骤S602:利用卷积神经网络,从获取到的多模态信息中提取图像特征;

[0099] 其中,所述卷积神经网络(Convolutional Neural Networks,CNN)可以是 是一类包含卷积计算且具有深度结构的前馈神经网络,所述卷积神经网络可以 是ResNet (Residual Network,残差网络),还可以是Fast R-CNN(Fast Region-based Convolution Neural Networks,基于区域的快速卷积神经网络)等。

[0100] 步骤S604:利用循环神经网络,从所述多模态信息中提取文本特征。

[0101] 其中,循环神经网络 (Recurrent Neural Network,RNN) 是一类以序列数据 为输入,在序列的演进方向进行递归且所有节点 (循环单元) 按链式连接的递 归神经网络;所述循环神经网络可以是GRU (Gate Recurrent Unit,门控循环单 元),还可以是LSTM (Long Short-Term Memory,长短期记忆网络) 等。

[0102] 其中,所述图像特征和所述文本特征均为至少一个;

[0103] 步骤S606:将至少一个图像特征和至少一个文本特征映射到第一空间维 度;

[0104] 针对每一所述文本特征,执行如下步骤:

[0105] 步骤S608:在所述第一空间维度下,基于注意力机制,确定每一所述图像 特征对应的第一特征向量和所述文本特征对应的第二特征向量;

[0106] 其中,在所述卷积神经网络为ResNet的情况下,所述注意力机制可以是 Top Down Attention机制;在所述卷积神经网络为Fast R-CNN的情况下,所述 注意力机制可以是 Bottom Up Attention机制。

[0107] 步骤S610:确定每一所述第一特征向量和所述第二特征向量之间的点积;

[0108] 步骤S612:根据每一所述点积,确定对应图像特征和所述文本特征之间的 相关度。

[0109] 步骤S614:根据每一所述相关度,确定对应图像特征的注意力权重;

[0110] 步骤S616:利用每一所述注意力权重,对各图像特征进行加权求和,得到 所述文本特征对应的加权图像特征。

[0111] 步骤S618:将每一所述加权图像特征和对应的所述文本特征映射到第二空 间维度;

[0112] 步骤S620:在所述第二空间维度下,确定所述加权图像特征对应的第三特 征向量和所述文本特征对应的第四特征向量;

[0113] 步骤S622:确定所述第三特征向量和所述第四特征向量之间的外积;

[0114] 步骤S624:对所述外积进行求和池化,得到所述加权图像特征对应的双模 态向量;

[0115] 步骤S626:根据所述双模态向量,生成所述多模态信息对应的意图分类结 果。

[0116] 本申请实施例中,不仅通过基于注意力机制和文本特征,对图像特征进行 注意力加权处理,得到加权图像特征,还对加权图像特征和文本特征进行多模 态分解双线性池化处理,得到双模态向量,进而生成意图分类结果,从而能够 使得图像特征和文本特征的每个点都进行了交互,融合效果更好。另外,分别 利用卷积神经网络,从获取到的多模态信息中提取图像特征;利用循环神经网 络,从所述多模态信息中提取文本特征,从而能够使得图像特征和文本特征的 提取更加准确。

[0117] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0118] 步骤S702:利用卷积神经网络,从获取到的多模态信息中提取图像特征;

[0119] 图4为本申请实施例提供的一种多模态信息的示意图,参见图4,假设本 申请实施例的多模态信息处理方法应用于客户401 (或称用户) 和商户 (或称 商家) 的对话场景中,客 户401在交互界面发来多模态信息,所述多模态信息 包含纯文本信息402“为什么没动静”, 以及一张订单及物流的截图403;则可 以利用卷积神经网络,从所述纯文本信息402和订单 及物流的截图403中提 取出图像特征,例如订单及物流的截图403中的鞋子的图片中包含

的颜色特征、纹理特征、形状特征和空间关系特征等。

[0120] 步骤S704:利用循环神经网络,从所述多模态信息中提取文本特征。

[0121] 其中,可以利用循环神经网络,从所述纯文本信息402中提取出文本特征“为何”、“什么”、“么没”、“没动”、“动静”等;还可以从订单及物流的截图 403中提取出文本特征“正在”、“出库”、“进入”、“第三方”、“卖家”、“仓库”、“准备”、“出库”、“催单”、“订单”、“物流”等。

[0122] 步骤S706:基于注意力机制,利用所述文本特征,对所述图像特征进行注意力加权处理,得到加权图像特征;

[0123] 步骤S708:对所述加权图像特征和所述文本特征进行多模态分解双线性池化处理,得到双模态向量;

[0124] 步骤S710:根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0125] 步骤S712:确定与所述意图分类结果匹配的知识点;

[0126] 步骤S714:根据所述知识点,生成多模态交互文本。

[0127] 本申请实施例中,通过根据与意图分类结果匹配的知识点,生成多模态交互文本,从而能够提高多模态交互文本的生成效率和准确度。

[0128] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:

[0129] 步骤S802:利用卷积神经网络,从获取到的多模态信息中提取图像特征;

[0130] 步骤S804:利用循环神经网络,从所述多模态信息中提取文本特征。

[0131] 步骤S806:基于注意力机制,利用所述文本特征,对所述图像特征进行注意力加权处理,得到加权图像特征;

[0132] 步骤S808:对所述加权图像特征和所述文本特征进行多模态分解双线性池化处理,得到双模态向量;

[0133] 步骤S810:根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0134] 步骤S812:确定与所述意图分类结果匹配的至少一个待选知识点;

[0135] 图5为本申请实施例提供的一种意图分类结果和知识点的匹配关系示意图,参见图5,可以根据所述多模态信息生成意图分类结果“订单及物流截图”501;可以将订单及物流截图501看作多模态信息的标签;可以先将所述订单及物流截图501映射到待选知识点,待选知识点是通过海量数据信息统计和人工标注得出的意图分类结果对应的最可能的回复知识点。所述待选知识点包括:查询物流5021、无物流更新5022、什么时间出库5023、退货物流5024、是否配送5025和无物流记录5026等。

[0136] 步骤S814:显示所述至少一个待选知识点;

[0137] 图6为本申请实施例提供的一种意图卡片的示意图,参见图6,可以根据待选知识点的使用频率在交互界面弹出意图卡片603,所述意图卡片603用于显示客户601可能的询问意图,所述意图卡片603上可以显示部分或全部的待选知识点。

[0138] 步骤S816:根据接收到的指令,从所述至少一个待选知识点中确定出目标知识点;

[0139] 其中,所述指令可以是根据所述用户601的点击操作或输入操作生成的待选知识点的选择指令;假设客户601点击或者输入意图卡片603中的“什么时间出库”,则商户操作的电子设备602可以将待选知识点“什么时间出库”确定为与目标知识点。

- [0140] 步骤S818:将所述目标知识点确定为与所述意图分类结果匹配的知识点。
- [0141] 步骤S820:根据所述知识点,生成多模态交互文本。
- [0142] 其中,电子设备可以根据客户601的点击操作,确定与对应待选知识点604 “什么时间出库”匹配的回复内容605,将所述回复内容作为多模态交互文本 显示在交互界面上,以供用户查看。所述待选知识点匹配的回复内容可由商家 根据自身实际情况配置,也可以电子设备根据待选知识点自动配置;在所述待 选知识点为“什么时间出库”的情况下,所述回复内容可以配置为“您好,非 常抱歉~因A地灾情严重,A地快件受到严重影响哦,具体发货时效无法保证 的呢”。
- [0143] 本申请实施例中,可以根据接收到的指令,从待选知识点中确定出目标知 识点,从而能够使得知识点的确定更加灵活。
- [0144] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:
- [0145] 步骤S902:利用卷积神经网络,从获取到的多模态信息中提取图像特征;
- [0146] 步骤S904:利用循环神经网络,从所述多模态信息中提取文本特征。
- [0147] 步骤S906:基于注意力机制,利用所述文本特征,对所述图像特征进行注 意力加权处理,得到加权图像特征;
- [0148] 步骤S908:对所述加权图像特征和所述文本特征进行多模态分解双线性池 化处理,得到双模态向量;
- [0149] 步骤S910:根据所述双模态向量,生成所述多模态信息对应的意图分类结 果。
- [0150] 步骤S912:确定与所述意图分类结果匹配的至少一个待选知识点;
- [0151] 步骤S914:确定每一所述待选知识点在历史时刻被确定为目标知识点的频 率;
- [0152] 其中,所述频率又称所述待选知识点的使用频率。
- [0153] 步骤S916:根据所述频率递减的顺序,对所述至少一个待选知识点进行排 序;
- [0154] 步骤S918:按照排列顺序,显示所述至少一个待选知识点。
- [0155] 其中,可以在所述意图卡片603上可以显示高频回复前三位的待选知识点: 什么时间出库、无物流记录、和无物流更新。
- [0156] 步骤S920:根据接收到的指令,从所述至少一个待选知识点中确定出目标 知识点;
- [0157] 其中,所述指令可以是根据所述用户601的点击操作或输入操作生成的待 选知识点的选择指令。
- [0158] 步骤S922:将所述目标知识点确定为与所述意图分类结果匹配的知识点。
- [0159] 步骤S924:根据所述知识点,生成多模态交互文本。
- [0160] 本申请实施例中,可以根据待选知识点的使用频率,确定所述待选知识点 的显示顺序,从而能够提升用户的交互体验。
- [0161] 本申请实施例还提供一种多模态信息处理方法,所述方法包括:
- [0162] 步骤S1002:利用卷积神经网络,从获取到的多模态信息中提取图像特征;
- [0163] 步骤S1004:利用循环神经网络,从所述多模态信息中提取文本特征。
- [0164] 步骤S1006:基于注意力机制,利用所述文本特征,对所述图像特征进行 注意力加权处理,得到加权图像特征;
- [0165] 步骤S1008:对所述加权图像特征和所述文本特征进行多模态分解双线性 池化处

理,得到双模态向量;

[0166] 步骤S1010:根据所述双模态向量,生成所述多模态信息对应的意图分类结果。

[0167] 步骤S1012:确定与所述意图分类结果匹配的至少一个待选知识点;

[0168] 步骤S1014:确定每一所述待选知识点在历史时刻被确定为目标知识点的频率;

[0169] 步骤S1016:将频率最高的待选知识点,确定为与所述意图分类结果匹配的知识点;

[0170] 其中,也可以直接将频率最高的待选知识点确定为与所述意图分类结果匹配的知识点。

[0171] 步骤S1018:根据所述知识点,生成多模态交互文本。

[0172] 本申请实施例中,还可以直接将频率最高的待选知识点确定为与所述意图分类结果匹配的知识点,从而提高了知识点确定的智能性。

[0173] 多模态应答是指计算机需要针对给定的包含文本和图像的信息进行智能应答。多模态应答可以称为多模态图像应答,其中,文本是一种模态,图像也是一种模态,同时包含图像和文本的信息被称之为多模态。与一般基于图像或者文本分类的应答方式不同,多模态涉及到两种或多种信息流,因此解决起来更困难,也更有挑战性。多模态信息应答在电商等智能客服中有极高的应用价值,因为在客服的会话中,用户不仅会发来纯文本信息,还可能包含图像信息,参见图4,用户发来了纯文本信息“为什么没动静”,以及图像信息,所述图像信息为订单及物流的截图。而传统的基于图像或文本分类的应答方式无法很好的解决此类多模态的问题,因此,通过多模态图像应答不仅可以节约人力成本,并且可以更加快速准确的回答用户的问题。

[0174] 相关技术中,多模态图像应答的解决方式主要有:基于单模态语言的分类或匹配模型的应答方式,如TextCNN(Text Convolutional Neural Networks,利用卷积神经网络进行文本分类)等;基于单模态图像分类模型的应答方式,如ResNet(Residual Network,残差网络)等;基于Attention(注意力)机制进行多模态融合的分类应答方式,通常是使用Attention的方式对两种信息进行融合,如SAN(Stacked Attention Networks,堆叠注意力网络)等模型,SAN是一种利用双层注意力机制的VQA(Visual Question Answering,视觉问答)网络。

[0175] 多模态应答中如果采用单模态(即只使用图像或者文本)方法进行解决,便浪费了其他模态的信息,得到的意图分类结果的准确率不高,而基于attention机制这类方法使用模态信息的交互,表现要比单模态的方法好一些,但是准确率仍然欠佳。

[0176] 同时,依托意图分类结果的回复需要一整套的后处理方法,仅仅回复意图分类结果会显得答案不够详细,不能很好解决的用户需求。

[0177] 单模态分类方法表现差的原因是由于多模态信息复杂的信息流,只是用一种信息无法充分理解上下文信息。参见图4,如果只使用文本信息而不对图像信息进行理解,客服仅根据文本信息“为什么没动静”,就无法理解用户是说的是“物流没动静”,就会由于信息理解不充分导致错误的应答。

[0178] 而通过Attention机制进行多模态信息交互,只对图像特征做了注意力加权,而文本信息中的文本特征只是在生成图像特征的注意力权重的过程中起作用,这样融合和交互的程度有限,可能导致文本中的一些重要信息被忽略,导致分类效果不佳。同时,依托

意图分类结果的回复需要一整套的处理方法,仅仅回复意图分类结果会显得答案不够详细,参见图4,如果只回复“订单及物流截图”,用户可能就不知所云,不能很好解决的用户问题,交互性也比较差。

[0179] 本申请实施例提出一种二段式交互方式的多模态应答方法,能够利用了多种模态信息,并使信息能够充分的交互融合,获得准确的多模态意图分类。同时为了使回复更加详细,本申请实施例采用基于二段式交互方式的“多模态意图分类-意图映射”回复方法,从而更好的进行应答。

[0180] 本申请实施例提供一种多模态应答方法,应用于电子设备,所述方法包括以下步骤:

[0181] 步骤S1102:通过多模态信息,生成意图分类结果。

[0182] 在一个应用场景中,用户咨询价保类的问题,在交互界面发送文本信息“为什么没动静”,以及图像信息:关于订单信息和物流信息的截图;通过本申请实施例提供的二段式多模态意图分类方法,可以得到意图分类结果为“订单及物流截图”。

[0183] 在整个多模态应答过程中,多模态意图识别是很最重要的一部分,因为用户的意图分类结果的准确性和接下来的应答质量直接相关。本申请实施例提出一种二段式的交互多模态意图分类模型来提高意图分类结果质量。

[0184] 图7为本申请实施例提供的一种意图分类结果的生成方法示意图,参见图7,首先,可以进行图像特征705和文本特征706的提取;可以通过卷积神经网络ResNet 703从图像信息701中提取多个图像特征705,通过双层循环神经网络GRU (Gate Recurrent Unit) 704从文本信息702中提取文本特征706。

[0185] 其次,在多模态交互的第一阶段,可以利用注意力机制(Attention) 707 进行多个图像特征和文本特征的融合,得到注意力加权的加权图像特征709。

[0186] 其中,通过Top down Attention(自上而下的注意力模型)对两种信息(即图像特征705和文本特征706)进行第一阶段的信息交互,即注意力阶段的融合,两种模态的信息首先被映射到一个共同的空间维度,接着进行点积融合,即计算每一图像特征和文本特征之间的点积,并通过softmax函数对每一所述点积进行归一化处理,得到每一图像特征的注意力权重708;根据每一图像的注意力权重708,对各图像特征进行加权求和,得到注意力加权的加权图像特征709。

[0187] 然而,此时的注意力机制只对图像特征进行加权求和,而文本信息在多模态对话中也是一个关键的信息,因此,接着,在多模态交互的第二阶段,对注意力加权的加权图像特征709和文本特征706进行MFB (Multi-modal Factorized Bilinear Pooling,多模态分解式双线性池)处理710得到多模态向量,并根据多模态向量,生成多模态信息对应的意图分类结果711。

[0188] 其中,MFB结构对于文本特征和注意力加权后的加权图像特征的每个点都进行了交互,融合效果更好,MFB结构如图3所示,分别将两种特征(注意力加权后的加权图像特征和文本特征)映射到更大的空间维度后,然后对注意力加权后的图像特征和文本特征进行求外积操作,最后作求和池化,这样能够让两种特征充分融合,并且最大程度保留了第一阶段被“忽视”的文本语义信息,更有利于多模态分类。

[0189] 步骤S1104:将所述意图分类结果映射到待选知识点。

[0190] 其中,待选知识点是通过海量数据信息统计和人工标注得出的分类对应的 最可能的回复知识点。参见图5,与“订单及物流截图”匹配的待选知识点包 括:查询物流、无物流更新、什么时间出库、退货物流、是否配送和无物流记 录。可以根据待选知识点的使用频率在交互界面弹出意图卡片,所述意图卡片 用于显示用户可能的询问意图,所述意图卡片上可以显示高频回复前三位的待 选知识点。用户点击意图卡片中的对应待选知识点,电子设备根据用户的点击 操作,确定与对应待选知识点匹配的回复内容,将所述回复内容作为多模态交 互文本显示在交互界面上,以供用户查看。其中,所述待选知识点匹配的回复 内容可由商家根据自身实际情况配置,在所述待选知识点为“什么时间出库” 的情况下,所述回复内容可以配置为“您好,非常抱歉~因A地灾情严重,A 地快件受到严重影响哦,具体发货时效无法保证的呢”。

[0191] 表1为本申请实施例提供的二段式交互方式的多模态应答方法,以及单模态 图像分类模型ResNet、堆叠注意力网络模型SAN、自上而下的注意力模型Top down Attention各自的分类准确率Acc、精确率Precision、召回率Recall以及精确 率和召回率的调和平均F1;

[0192] 表1

模型	Acc	Precision	召回率	F1
ResNet	0.901	0.769	0.730	0.749
SAN	0.889	0.747	0.700	0.716
Top down Attention	0.903	0.762	0.736	0.750
二段式交互方式	0.927	0.766	0.782	0.766

[0194] 参见表1,本申请实施例提供的二段式交互方式的多模态应答方法在最重 要的两个指标Acc (准确率)和F1 (精确率和召回率的调和平均) 都达到最优,从而提高了确定意图分类结果的准确性。

[0195] 基于前述的实施例,本申请实施例提供一种多模态信息处理装置,该装置 包括所包括的各模块,可以通过电子设备中的处理器来实现;当然也可通过具 体的逻辑电路实现;在实施的过程中,处理器可以为中央处理器(CPU,Central Processing Unit)、微处理器(MPU, Microprocessor Unit)、数字信号处理器(DSP, Digital Signal Processing)或现场可编程门阵列(FPGA,Field Programmable Gate Array)等。

[0196] 图8为本申请实施例多模态信息处理装置的组成结构示意图,如图8所示, 所述装置800包括提取模块801、加权模块802、池化模块803和第一生成模块 804,其中:

[0197] 提取模块801,用于从获取到的多模态信息中提取第一模态特征和第二模 态特征;所述第一模态特征和所述第二模态特征为两种不同模态的特征;

[0198] 加权模块802,用于基于注意力机制,利用所述第二模态特征,对所述第 一模态特征进行注意力加权处理,得到第一加权模态特征;

[0199] 池化模块803,用于对所述第一加权模态特征和所述第二模态特征进行多 模态分解双线性池化处理,得到双模态向量;

[0200] 第一生成模块804,用于根据所述双模态向量,生成所述多模态信息对应 的意图分类结果。

[0201] 在一个实施例中,所述第一模态特征和所述第二模态特征均为至少一个时,所述

加权模块802,包括:第一映射子模块,用于将至少一个第一模态特征和至少一个第二模态特征映射到第一空间维度;第一确定子模块,用于针对每一所述第二模态特征,在所述第一空间维度下,基于注意力机制,确定每一所述第一模态特征和所述第二模态特征之间的相关度;第二确定子模块,用于针对每一所述第二模态特征,根据每一所述相关度,确定对应第一模态特征的注意力权重;加权子模块,用于针对每一所述第二模态特征,利用每一所述注意力权重,对各第一模态特征进行加权求和,得到所述第二模态特征对应的第一加权模态特征。

[0202] 在一个实施例中,所述第一确定子模块,包括:第一确定单元,用于针对每一所述第二模态特征,确定每一所述第一模态特征对应的第一特征向量和所述第二模态特征对应的第二特征向量;第二确定单元,用于针对每一所述第二模态特征,确定每一所述第一特征向量和所述第二特征向量之间的点积;第三确定单元,用于针对每一所述第二模态特征,根据每一所述点积,确定对应第一模态特征和所述第二模态特征之间的相关度。

[0203] 在一个实施例中,所述第一加权模态特征和所述第二模态特征均为至少一个时,所述池化模块803,包括:第二映射子模块,用于将每一所述第一加权模态特征和对应的所述第二模态特征映射到第二空间维度;第三确定子模块,用于在所述第二空间维度下,确定所述第一加权模态特征对应的第三特征向量和所述第二模态特征对应的第四特征向量;第四确定子模块,用于确定所述第三特征向量和所述第四特征向量之间的外积;池化子模块,用于对所述外积进行求和池化,得到所述第一加权模态特征对应的双模态向量。

[0204] 在一个实施例中,所述第一模态特征包括图像特征,所述第二模态信息包括文本特征,所述提取模块,包括:第一提取子模块,用于利用卷积神经网络,从获取到的多模态信息中提取图像特征;第二提取子模块,用于利用循环神经网络,从所述多模态信息中提取文本特征。

[0205] 在一个实施例中,所述装置还包括:确定模块,用于确定与所述意图分类结果匹配的知识点;第二生成模块,用于根据所述知识点,生成多模态交互文本。

[0206] 在一个实施例中,所述确定模块,包括:第五确定子模块,用于确定与所述意图分类结果匹配的至少一个待选知识点;显示子模块,用于显示所述至少一个待选知识点;第六确定子模块,用于根据接收到的指令,从所述至少一个待选知识点中确定出目标知识点;第七确定子模块,用于将所述目标知识点确定为与所述意图分类结果匹配的知识点。

[0207] 在一个实施例中,所述显示子模块,包括:第四确定单元,用于确定每一所述待选知识点在历史时刻被确定为目标知识点的频率;排序单元,用于根据所述频率递减的顺序,对所述至少一个待选知识点进行排序;显示单元,用于按照排列顺序,显示所述至少一个待选知识点。

[0208] 在一个实施例中,所述确定模块,包括:第八确定子模块,用于确定与所述意图分类结果匹配的至少一个待选知识点;第九确定子模块,用于确定每一所述待选知识点在历史时刻被确定为目标知识点的频率;第十确定子模块,用于将频率最高的待选知识点,确定为与所述意图分类结果匹配的知识点。

[0209] 需要说明的是,本申请实施例中,如果以软件功能模块的形式实现上述的多模态信息处理方法,并作为独立的产品销售或使用,也可以存储在一个计算机可读存储介质中。基于这样的理解,本申请实施例的技术方案本质上或者说对相关技术做出贡献的部

分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得电子设备(可以是手机、平板电脑、台式机、个人数字助理、数字电话、视频电话、电视机、传感设备等)执行本申请各个实施例所述方法的全部或部分。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read Only Memory,ROM)、磁碟或者光盘等各种可以存储程序代码的介质。这样,本申请实施例不限制于任何特定的硬件和软件结合。

[0210] 以上装置实施例的描述,与上述方法实施例的描述是类似的,具有同方法实施例相似的有益效果。对于本申请装置实施例中未披露的技术细节,请参照本申请方法实施例的描述而理解。

[0211] 对应地,本申请实施例提供一种电子设备,图9为本申请实施例电子设备的一种硬件实体示意图,如图9所示,该电子设备900的硬件实体包括:包括存储器901和处理器902,所述存储器901存储有可在处理器902上运行的计算机程序,所述处理器902执行所述程序时实现上述实施例多模态信息处理方法中的步骤。

[0212] 存储器901配置为存储由处理器902可执行的指令和应用,还可以缓存待处理器902以及电子设备900中各模块待处理或已经处理的数据(例如,图像数据、音频数据、语音通信数据和视频通信数据),可以通过闪存(FLASH)或随机访问存储器(Random Access Memory,RAM)实现。

[0213] 对应地,本申请实施例提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现上述实施例中提供的多模态信息处理方法中的步骤。

[0214] 这里需要指出的是:以上存储介质和设备实施例的描述,与上述方法实施例的描述是类似的,具有同设备实施例相似的有益效果。对于本申请存储介质和方法实施例中未披露的技术细节,请参照本申请设备实施例的描述而理解。

[0215] 应理解,说明书通篇中提到的“一个实施例”或“一实施例”意味着与实施例有关的特定特征、结构或特性包括在本申请的至少一个实施例中。因此,在整个说明书各处出现的“在一个实施例中”或“在一实施例中”未必一定指相同的实施例。此外,这些特定的特征、结构或特性可以任意适合的方式结合在一个或多个实施例中。应理解,在本申请的各种实施例中,上述各过程的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不对本申请实施例的实施过程构成任何限定。上述本申请实施例序号仅仅为了描述,不代表实施例的优劣。

[0216] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者装置不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者装置所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。

[0217] 在本申请所提供的几个实施例中,应该理解到,所揭露的设备和方法,可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的,例如,所述单元的划分,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,如:多个单元或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口,设备或单元的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0218] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的,作为单元显示的部件可以是、或也可以不是物理单元;既可以位于一个地方,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全部单元来实现本实施例方案的目的。另外,在本申请各实施例中的各功能单元可以全部集成在一个处理单元中,也可以是各单元分别单独作为一个单元,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,也可以采用硬件加软件功能单元的形式实现。

[0219] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(Read Only Memory,ROM)、磁碟或者光盘等各种可以存储程序代码的介质。或者,本申请上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用,也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,本申请实施例的技术方案本质上或者说对相关技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得电子设备(可以是手机、平板电脑、台式机、个人数字助理、数字电话、视频电话、电视机、传感设备等)执行本申请各个实施例所述方法的全部或部分。而前述的存储介质包括:移动存储设备、ROM、磁碟或者光盘等各种可以存储程序代码的介质。

[0220] 本申请所提供的几个方法实施例中所揭露的方法,在不冲突的情况下可以任意组合,得到新的方法实施例。本申请所提供的几个产品实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的产品实施例。本申请所提供的几个方法或设备实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的方法实施例或设备实施例。

[0221] 以上所述,仅为本申请的实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

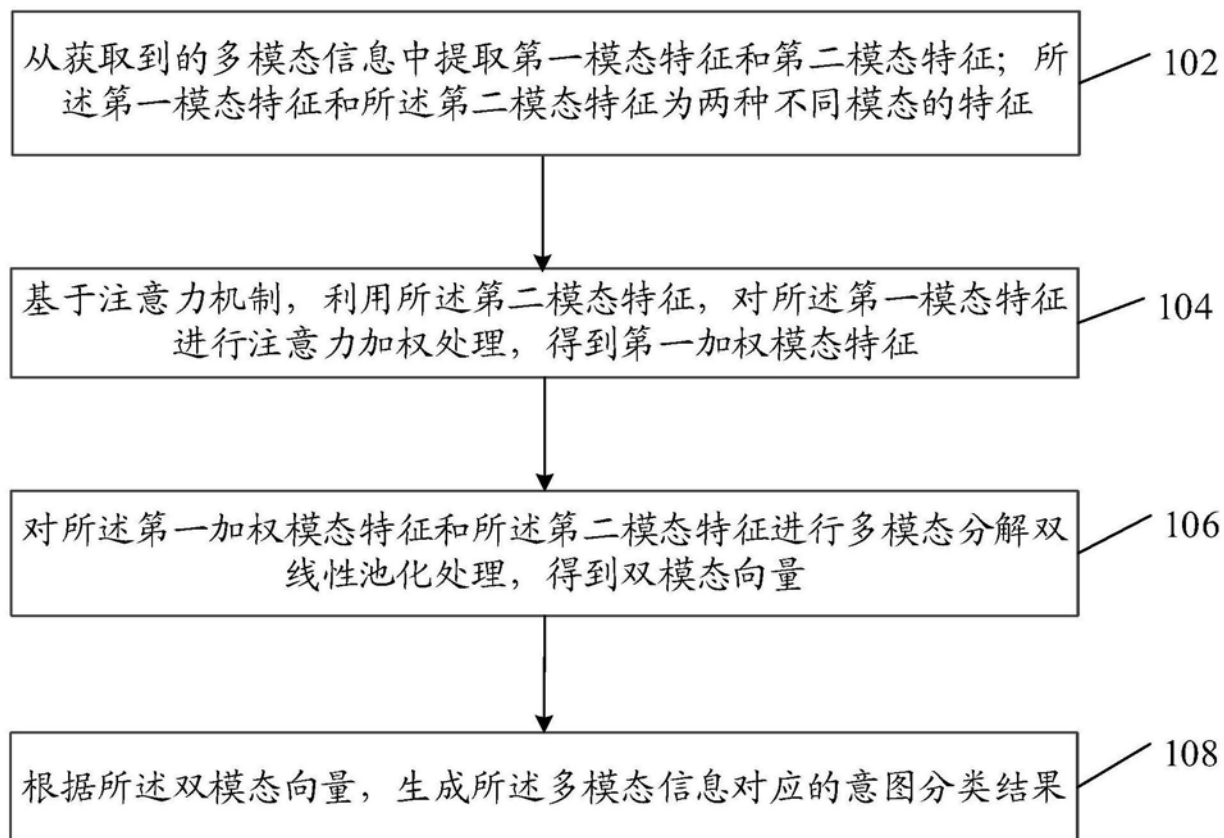


图1

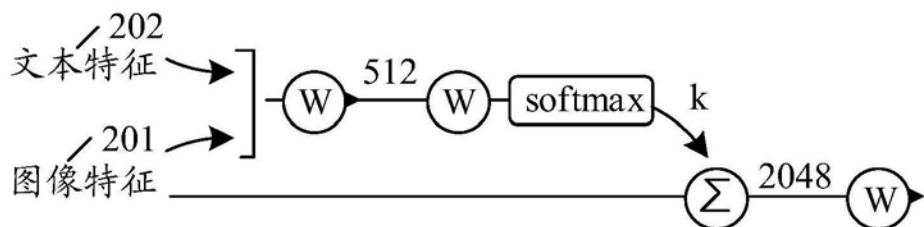


图2

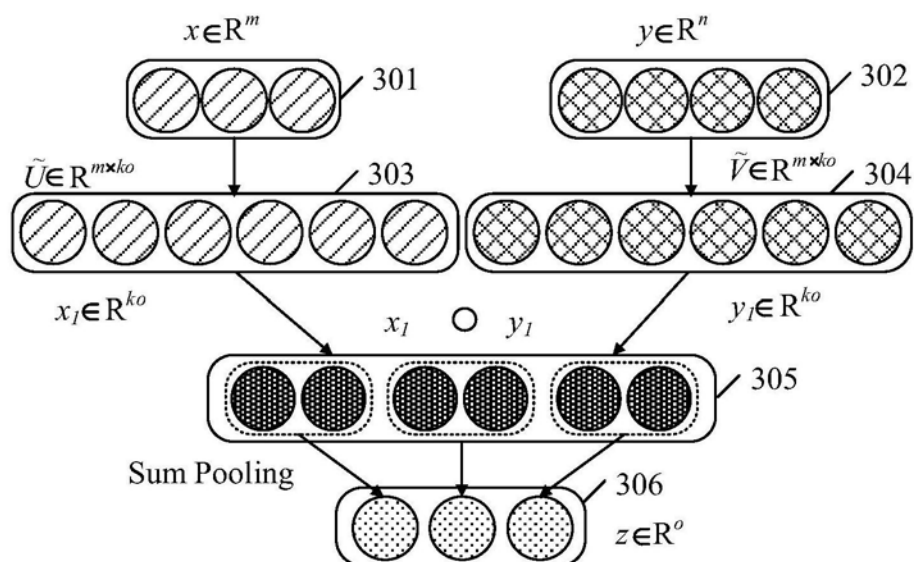


图3

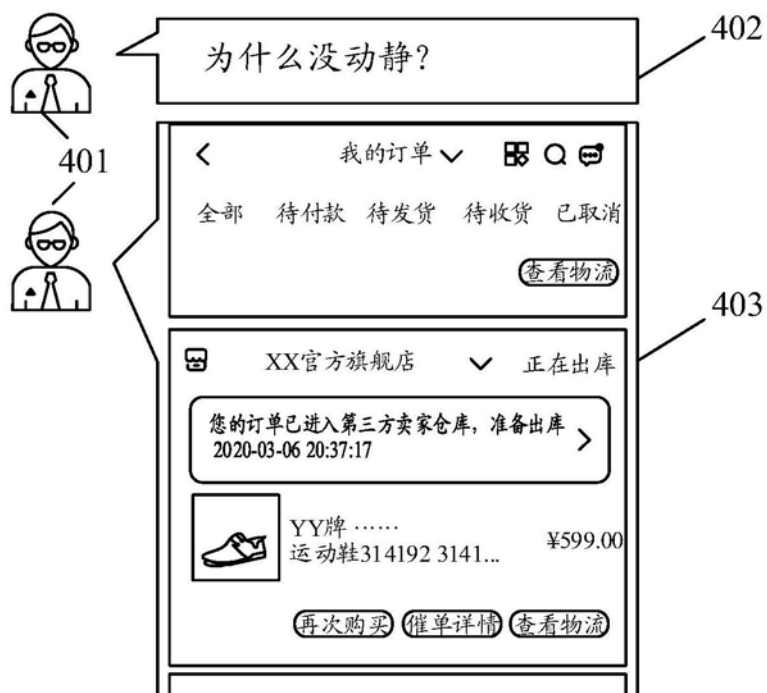


图4

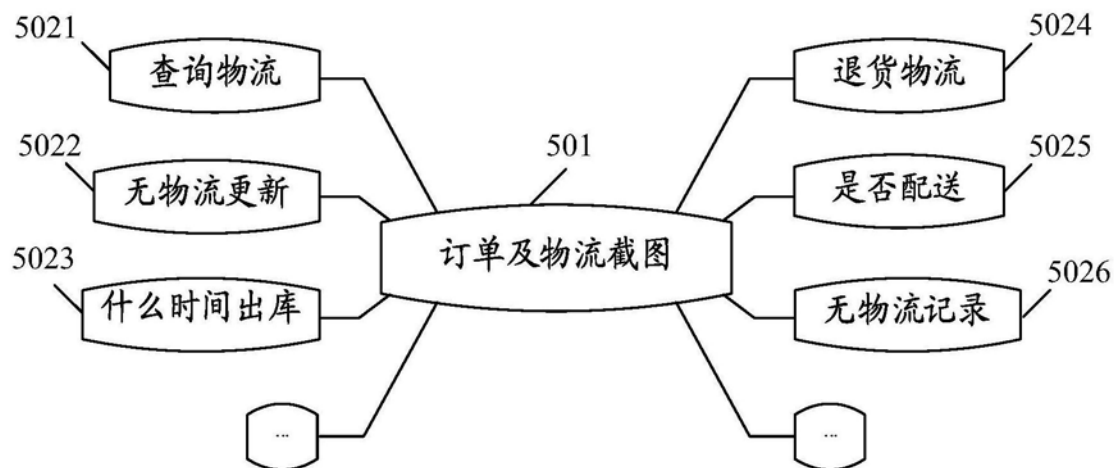


图5

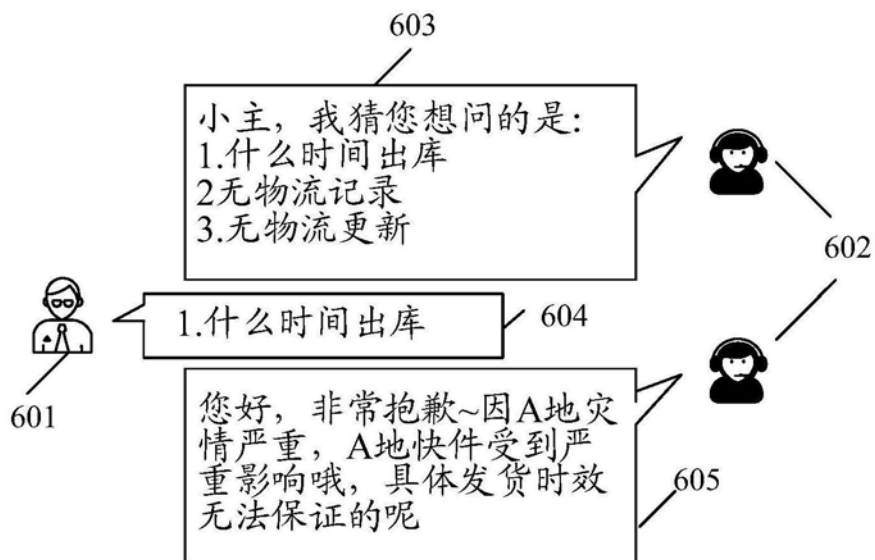


图6

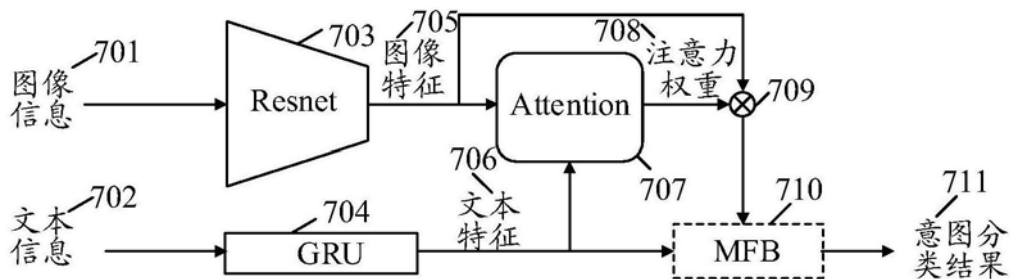


图7

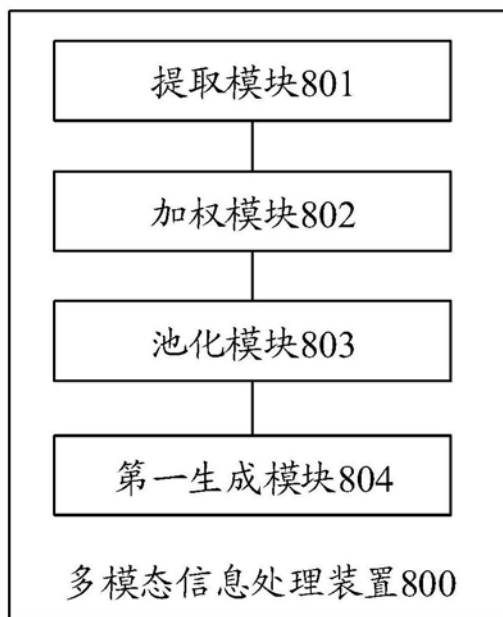


图8

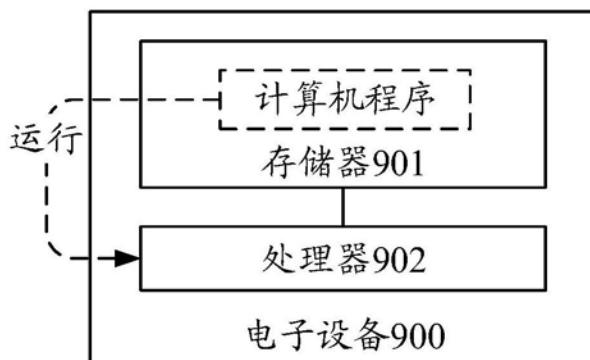


图9