



(12)发明专利申请

(10)申请公布号 CN 111680159 A

(43)申请公布日 2020.09.18

(21)申请号 202010532007.3

(22)申请日 2020.06.11

(71)申请人 华东交通大学

地址 330013 江西省南昌市青山湖区双港  
东路808号

(72)发明人 姜楠 李进 黄伟 陈鸿龙

(74)专利代理机构 深圳市智圈知识产权代理事  
务所(普通合伙) 44351

代理人 吕静

(51)Int.Cl.

G06F 16/35(2019.01)

G06F 40/35(2020.01)

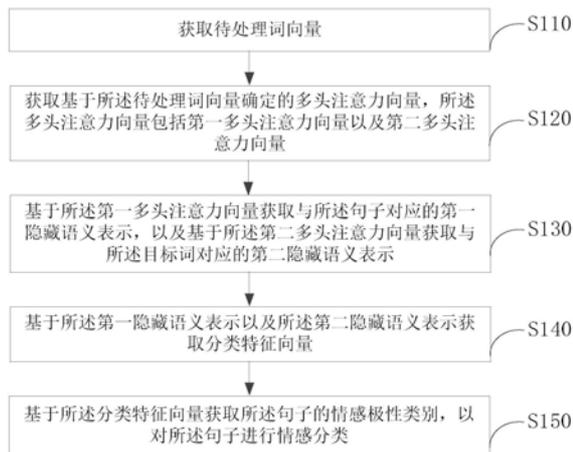
权利要求书2页 说明书11页 附图4页

(54)发明名称

数据处理方法、装置及电子设备

(57)摘要

本申请公开一种数据处理方法、装置、电子设备及存储介质。该方法包括：获取待处理词向量，继而获取基于所述待处理词向量确定的多头注意力向量，再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示，以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示，继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量，然后基于所述分类特征向量获取所述句子的情感极性类别，以对所述句子进行情感分类。实现了通过获取基于待处理词向量确定的多头注意力向量，以获得与句子和目标词对应的融合了句子的全句语义的向量表示，从而使得分类特征向量的分类结果更加准确，进而可以提升情感分类的准确度。



1. 一种数据处理方法,其特征在于,所述方法包括:

获取待处理词向量,所述待处理词向量包括与句子和目标词分别对应的词向量、以及与所述句子和所述目标词分别对应的参考词向量;

获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;

基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;

基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;

基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。

2. 根据权利要求1所述的方法,其特征在于,所述获取待处理词向量,包括:

获取待处理句子和待处理目标词,所述待处理句子包括所述待处理目标词;

将所述待处理句子和所述待处理目标词输入BERT模型,所述BERT模型为预训练得到,获取所述BERT模型输出的与所述待处理句子和所述待处理目标词分别对应的向量表示,所述向量表示中所述待处理句子和所述待处理目标词对应,所述向量表示融合了全句语义信息;

对所述向量表示进行求平均操作,得到待处理词向量。

3. 根据权利要求1所述的方法,其特征在于,所述获取基于所述待处理词向量确定的多头注意力向量,包括:

将所述待处理词向量输入目标神经网络的多头注意力网络层,得到所述多头注意力网络层输出的多头注意力向量。

4. 根据权利要求3所述的方法,其特征在于,所述将所述待处理词向量输入目标神经网络的多头注意力网络层,得到所述多头注意力网络层输出的多头注意力向量,包括:

将与所述句子对应的词向量和与所述目标词对应的参考词向量输入多头注意力函数,输出得到第一多头注意力向量;

将与所述目标词对应的词向量和与所述句子对应的参考词向量输入多头注意力函数,输出得到第二多头注意力向量。

5. 根据权利要求4所述的方法,其特征在于,所述目标神经网络还包括卷积层,所述基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示,包括:

将所述第一多头注意力向量输入所述卷积层,得到第一隐藏状态;

将所述第二多头注意力向量输入所述卷积层,得到第二隐藏状态;

对所述第一隐藏状态进行求平均操作,得到与所述句子对应的第一隐藏语义表示;

对所述第二隐藏状态进行求平均操作,得到与所述目标词对应的第二隐藏语义表示。

6. 根据权利要求5所述的方法,其特征在于,所述基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量,包括:

基于预设规则将所述第一隐藏语义表示与所述第二隐藏语义表示进行连接,得到分类特征向量。

7. 根据权利要求1所述的方法,其特征在于,所述基于所述分类特征向量获取所述句子的情感极性类别,包括:

对所述分类特征向量进行线性特征变换,将线性变换后的结果投影到指定空间,获取所述指定空间下所述句子对应的情感极性类别的预测概率;

将预测概率最大对应的情感极性类别作为所述句子的情感极性类别。

8. 一种数据处理装置,其特征在于,所述装置包括:

词向量获取模块,用于获取待处理词向量,所述待处理词向量包括与句子和目标词分别对应的词向量、以及与所述句子和所述目标词分别对应的参考词向量;

多头注意力向量获取模块,用于获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;

隐藏语义表示获取模块,用于基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;

分类特征向量获取模块,用于基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;

情感极性类别获取模块,用于基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。

9. 一种电子设备,其特征在于,包括一个或多个处理器以及存储器;

一个或多个程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序配置用于执行权利要求1-7任一所述的方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有程序代码,其中,在所述程序代码被处理器运行时执行权利要求1-7任一所述的方法。

## 数据处理方法、装置及电子设备

### 技术领域

[0001] 本申请涉及自然语言处理技术领域,更具体地,涉及一种数据处理方法、装置、电子设备及存储介质。

### 背景技术

[0002] 随着互联网技术的迅猛发展与广泛应用,越来越多的用户倾向于通过网络的方式发表意见、表达情绪或阐述观点。情感分析,又称为意见挖掘,是分析人们对于产品、服务、组织、个人、事件、主题及其属性等实体对象所怀有的意见、情感、评价、看法和态度等主观感受的研究领域。作为一种方式,可以通过情感分析的方式挖掘用户的意见、情绪或观点等,通过情感分析,可以帮助用户高效、准确以及全面地从海量文本信息中获取人们所关注的内容。然而,目前在对包含多种情感或多个方面的文本进行情感分析时,对同一文本中不同方面的情感极性的区分度以及情感分类的准确度有待提升。

### 发明内容

[0003] 鉴于上述问题,本申请提出了一种数据处理方法、装置、电子设备及存储介质,以改善上述问题。

[0004] 第一方面,本申请实施例提供了一种数据处理方法,该方法包括:获取待处理词向量,所述待处理词向量包括与句子和目标词分别对应的词向量、以及与所述句子和所述目标词分别对应的参考词向量;获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。

[0005] 第二方面,本申请实施例提供了一种数据处理装置,该装置包括:词向量获取模块,用于获取待处理词向量,所述待处理词向量包括与句子和目标词分别对应的词向量、以及与所述句子和所述目标词分别对应的参考词向量;多头注意力向量获取模块,用于获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;隐藏语义表示获取模块,用于基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;分类特征向量获取模块,用于基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;情感极性类别获取模块,用于基于所述分类特征向

量获取所述句子的情感极性类别,以对所述句子进行情感分类。

[0006] 第三方面,本申请实施例提供了一种电子设备,包括存储器以及一个或多个处理器;一个或多个程序被存储在存储器中并被配置为由一个或多个处理器执行,一个或多个程序配置用于执行上述第一方面所述的方法。

[0007] 第四方面,本申请实施例提供了一种计算机可读存储介质,计算机可读存储介质中存储有程序代码,其中,在程序代码由处理器运行时执行上述第一方面所述的方法。

[0008] 本申请实施例提供一种数据处理方法、装置、电子设备以及存储介质,通过获取待处理词向量,继而获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;然后基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。从而通过上述方式实现了通过获取基于待处理词向量确定的多头注意力向量,以获得与句子和目标词对应的融合了句子的全句语义的向量表示,从而使得分类特征向量的分类结果更加准确,进而可以提升情感分类的准确度。

## 附图说明

[0009] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述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它的附图。

[0010] 图1示出了本申请实施例提供的一种数据处理方法的方法流程图。

[0011] 图2示出了图1中的方法S110的方法流程图。

[0012] 图3示出了本申请实施例中的目标神经网络对句子进行方面级文本情感分类的处理过程示例图。

[0013] 图4示出了本申请实施例提供的一种数据处理装置的结构框图。

[0014] 图5示出了本申请实施例提供的一种电子设备的结构框图。

[0015] 图6示出了本申请实施例的用于保存或者携带实现根据本申请实施例的数据处理方法的程序代码的存储单元。

## 具体实施方式

[0016] 为了使本技术领域的人员更好地理解本申请方案,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。

[0017] 情感分析是自然语言处理中的一项基本任务,可以用来挖掘用户的意见,进行数据分析、舆情监控等任务。情感分析又可分为情感信息的抽取、情感信息的分类以及情感信息的检索与归纳。

[0018] 其中,情感分类是指根据文本所表达的含义和情感信息将文本划分成褒扬的或贬

义的两种或几种类型,是对文本作者倾向性和观点、态度的划分。情感分类包括三种不同级别,分别是文档级别、句子级别和方面级别 (Aspect Level)。其中,文档级情感分类将观点鲜明的文档(例如,产品评论)分类为整体积极的或消极的观点。它将整个文档当做基本的信息单元,并假定文档是观点鲜明的,包含对单个实体(例如,某个型号的手机)的观点。句子级情感分类对文档内单独的语句进行分类。然而,单独的句子不能假定为观点鲜明的。和文档级、语句级情感分类相比,方面级别的情感分类更加细粒化。它的任务是提取和总结人们对某实体的观点以及实体(也被称为目标、方面特征词 (Aspect)) 的特征。例如一篇产品评论,方面级别的情感分类的目的是分别总结对产品或事件不同方面的积极和消极观点。

[0019] 当一个句子中存在多个方面特征词的情况下,为了分析不同方面特征词的情感极性,我们需要一个特定的模型来完成任务,这与传统的情感分类模型不同。例如,在评论句子“餐厅的食物味道很好,但是服务员的服务态度很差”中,有两个方面特征词:“味道”和“服务态度”,其中“味道”的情感极性是积极的,而“服务态度”的情感极性是消极的。该例子的整个句子的情感极性由积极和消极两种组成。如果我们不考虑方面特征词的信息,就很难判断这个句子的情感极性,这种类型的错误通常存在于一般的情感分类任务中。

[0020] 目前,一般基于AOA模型处理情感分类任务,该模型使用了双重注意力机制,具体处理步骤如下:

[0021] 第一步,将句子和目标词的词向量分别通过双向长短期记忆网络 (BiLSTM) 后得到的隐藏状态矩阵进行相乘操作得到矩阵I,计算公式如下:

$$[0022] \quad I = h_s \cdot h_t^T。$$

[0023] 第二步,通过Softmax函数计算两个注意力矩阵 $\alpha_{ij}$ 和 $\beta_{ij}$ :

$$[0024] \quad \alpha_{ij} = \frac{\exp(I_{ij})}{\sum_i \exp(I_{ij})},$$

$$[0025] \quad \beta_{ij} = \frac{\exp(I_{ij})}{\sum_j \exp(I_{ij})}。$$

[0026] 第三步,对注意力矩阵 $\beta_{ij}$ 求平均,得到目标词级别的注意力向量 $\bar{\beta}$ :

$$[0027] \quad \bar{\beta} = \sum_i \beta_{ij} / n。$$

[0028] 第四步,将注意力矩阵 $\alpha_{ij}$ 与注意力向量 $\bar{\beta}$ 相乘,得到最终的句子级注意力向量 $\gamma$ :

$$[0029] \quad \gamma = \alpha \cdot \bar{\beta}^T。$$

[0030] 第五步,将句子的隐藏状态矩阵与最终的句子级注意力向量 $\gamma$ 相乘得到最终的句子表示r:

$$[0031] \quad r = h_s^T \cdot \gamma。$$

[0032] 第六步,将最终的句子表示r输入线性函数后,通过Softmax函数进行分类:

$$[0033] \quad x = W_1 \cdot r + b_1,$$

$$[0034] \quad P(y=c) = \frac{\exp(x_c)}{\sum_{i \in C} \exp(x_i)}。$$

[0035] 其中,  $W_1$  表示权重矩阵,  $b_1$  表示偏置项,  $C$  表示情感极性的类别数。

[0036] 而上述方法忽略了句子与方面特征词之间的相互依赖性, 在区分一个句子中不同方面的情感极性和提高情感分类的准确度方面仍有待改进, 并且该模型无法生成输入的句子和目标词对应的融合全句语义信息的向量表示。

[0037] 因此, 为了改善上述问题, 本申请实施例提供了一种数据处理方法、装置、电子设备以及存储介质。本申请实施例中的数据处理方法可以应用于分析产品评论或者是服务评论等场景。本方法通过获取待处理词向量, 继而获取基于所述待处理词向量确定的多头注意力向量, 所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量, 所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示, 所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示; 再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示, 以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示; 继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量; 然后基于所述分类特征向量获取所述句子的情感极性类别, 以对所述句子进行情感分类。从而通过上述方式实现了通过获取基于待处理词向量确定的多头注意力向量, 以获得与句子和目标词对应的融合了句子的全句语义的向量表示, 从而使得分类特征向量的分类结果更加准确, 进而可以提升情感分类的准确度。

[0038] 下面将结合附图具体描述本申请的各实施例。

[0039] 请参阅图1, 示出了本申请一实施例提供的一种数据处理方法的流程图, 本实施例提供一种数据处理方法, 可应用于电子设备, 所述电子设备可以是智能手机、平板、电脑、可穿戴电子设备、服务器等可运行程序的设备。该方法包括:

[0040] 步骤S110: 获取待处理词向量。

[0041] 本实施例中的待处理词向量可以包括与句子和目标词分别对应的词向量、以及与句子和目标词分别对应的参考词向量。其中, 与句子对应的词向量可以表示为  $e_s$ , 与目标词对应的词向量可以表示为  $e_t$ 。与句子和目标词分别对应的参考词向量可以理解为输入的句子和目标词对应、且融合了全句语义信息的向量表示, 其中, 与句子对应的参考词向量可以表示为  $\bar{e}_s$ , 与目标词对应的参考词向量可以表示为  $\bar{e}_t$ 。

[0042] 可选的, 待处理向量的具体获取过程描述如下:

[0043] 请参阅图2, 作为一种方式, 步骤S110可以包括:

[0044] 步骤S111: 获取待处理句子和待处理目标词, 所述待处理句子包括所述待处理目标词。

[0045] 其中, 待处理句子可以是长度为  $m$  的句子  $s = [w_1, w_2, \dots, w_i, \dots, w_m]$  (其中,  $w$  可以表示长度为  $m$  的句子中的词语或者单词, 数字“ $1, 2, \dots, i, \dots, m$ ”可以表示句子中的第几个词语或者单词)。待处理目标词可以是长度为  $n$  的目标词  $t = [w_i, w_{i+1}, \dots, w_{i+n-1}]$ , 可选的,  $w$  可以表示长度为  $m$  的句子中的词语或者单词, 数字“ $i, i+1, \dots, i+n-1$ ”可以表示句子中的第几个词语或者单词。需要说明的是, 本实施例中的待处理句子包括待处理目标词, 即待处理目标词是来源于待处理句子中的目标词, 例如, 可以是待处理句子中提取的词语。

[0046] 可选的,待处理目标词可以是一个句子中词性为名词的词语,具体的,本实施例中的待处理目标词可以是方面特征词。例如在句子“餐厅的食物味道很好,但是服务员的服务态度很差”中,方面特征词可以包括“味道”、“服务态度”,那么待处理目标词为:“味道”、“服务态度”。

[0047] 作为一种方式,可以通过人工标注的方式获取待处理句子,即通过对句子中的词语进行标注的方式获取待处理句子。在这种方式下,可以基于特征词提取模型从已获取的句子中提取方面特征词,其中,特征词提取模型可基于斯坦福核心自然语言处理数据集(Stanford CoreNLP)训练得到,斯坦福核心数据集是一个集成多种工具的自然语言处理(Natural Language Processing,NLP)工具集,工具包括词性标注(Part of Speech Tagger),包括标注有各单词词性的文本数据。利用训练得到的特征词提取模型可以从句子中提取方面特征词,进而获得待处理目标词。

[0048] 步骤S112:将所述待处理句子和所述待处理目标词输入BERT模型,获取所述BERT模型输出的与所述待处理句子和所述待处理目标词分别对应的向量表示。

[0049] 其中,BERT模型(Bidirectional Encoder Representations from Transformers,来自变换器的双向编码器表征量)为预训练得到。可选的,向量表示中待处理句子和待处理目标词对应,向量表示融合有全句语义信息。

[0050] 作为一种方式,在获取了待处理句子和待处理目标词后,可以将待处理句子和待处理目标词输入BERT模型,进而可以获取BERT模型输出的与待处理句子和待处理目标词分别对应的向量表示。具体的,将前述待处理句子 $s = [w_1, w_2, \dots, w_i, \dots, w_m]$ 输入到BERT模型后,可以输出得到与待处理句子对应的向量表示为: $e^s = [e_s^1; e_s^2; \dots; e_s^i; \dots; e_s^m]$ ;将前述待处理目标词 $t = [w_i, w_{i+1}, \dots, w_{i+n-1}]$ 输入到BERT模型后,可以输出得到与待处理目标词对应的向量表示为: $e^t = [e_T^i; e_T^{i+1}; \dots; e_T^{i+n-1}]$ 。

[0051] 步骤S113:对所述向量表示进行求平均操作,得到待处理词向量。

[0052] 作为一种方式,在通过BERT模型获取到了与待处理句子和待处理目标词分别对应的向量表示后,可以分别对二者进行求平均操作,以获取待处理词向量。其中,通过对与待处理句子对应的向量表示 $e^s = [e_s^1; e_s^2; \dots; e_s^i; \dots; e_s^m]$ 求平均操作,可以得到:

$$[0053] \quad \bar{e}_s = \sum_{i=1}^m \frac{e_s^i}{m},$$

[0054] 通过对与待处理目标词对应的向量表示 $e^t = [e_T^i; e_T^{i+1}; \dots; e_T^{i+n-1}]$ 求平均操作。可以得到:

$$[0055] \quad \bar{e}_T = \sum_{i=1}^n \frac{e_T^i}{n}。$$

[0056] 可选的,可以将得到的与待处理句子对应的向量表示和与待处理目标词对应的向量表示作为待处理词向量。

[0057] 步骤S120:获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量。

[0058] 作为一种方式,可以将待处理词向量输入目标神经网络的多头注意力网络层(简

称为MHA),以获取多头注意力网络层输出的多头注意力向量,其中,多头注意力向量可以包括第一多头注意力向量以及第二多头注意力向量。可选的,第一多头注意力向量可以用于表征句子对目标词的多头注意力向量表示,第二多头注意力向量可以用于表征目标词对句子的多头注意力向量表示。

[0059] 其中,目标神经网络为基于共同注意力机制的情感分类神经网络模型,可选的,目标神经网络可以包括多头注意力网络层和卷积层。可选的,多头注意力网络层的工作原理描述如下:

[0060] 假设有一组键值对 $source = \langle key, value \rangle$ 和一组查询 $query$ ,针对每一次查询操作可以计算键 $key$ 与查询 $query$ 之间的相似度。可选的,可以定义查询 $query$ 向量为: $q = [q_1, q_2, \dots, q_u]$ ,键 $key$ 向量为: $k = [k_1, k_2, \dots, k_v]$ ,在这种方式下,注意力值的计算公式可以为:

[0061]  $f_s(k, q) = \tanh([k_i; q_j] \cdot W_a), i = 1, 2, \dots, u; j = 1, 2, \dots, v$ 。

[0062] 可选的,上式中 $W_a$ 可以表示权重矩阵。

[0063] 然后,可以将注意力值 $f_s(k, q)$ 通过softmax函数后求和,计算注意力值 $Attention(k, q)$ :

[0064]  $Attention(k, q) = \sum_1^n \text{softmax}(f_s(k, q))k$ ,

[0065] 可选的,可以采用多头注意力并行地计算输入信息,计算公式如下:

[0066]  $head_i = Attention(k, q)$ ,

[0067]  $MHA(k, q) = (head_1; head_2; \dots; head_h) \cdot W_o$ 。

[0068] 可选的,上式中 $head_i$ 表示第 $i$ 个头注意力, $W_o$ 为权重矩阵。

[0069] 基于上述描述,作为一种方式,可以将与句子对应的词向量和与目标词对应的参考词向量输入多头注意力函数,即将 $e_s$ 和 $\bar{e}_T$ 输入函数 $MHA(k, q) = (head_1; head_2; \dots; head_h) \cdot W_o$ ,进而输出得到第一多头注意力向量:

[0070]  $s^a = MHA(e_s, \bar{e}_T)$ 。

[0071] 可选的,在将 $e_s$ 和 $\bar{e}_T$ 通过多头注意力网络层的时候,可以将句子对目标词的多头注意力表示为: $s^a = [s_1^a; s_2^a; \dots; s_m^a]$ 。

[0072] 类似的,可以将与目标词对应的词向量和与句子对应的参考词向量输入多头注意力函数,即将 $e_T$ 和 $\bar{e}_S$ 输入函数 $MHA(k, q) = (head_1; head_2; \dots; head_h) \cdot W_o$ ,进而输出得到第二多头注意力向量:

[0073]  $t^b = MHA(e_T, \bar{e}_S)$ 。

[0074] 可选的,在将 $e_T$ 和 $\bar{e}_S$ 通过多头注意力网络层的时候,可以将目标词对句子的多头注意力表示为: $t^b = [t_i^b; t_{i+1}^b; \dots; t_{i+n-1}^b]$ 。

[0075] 可选的,上述步骤可以通过如下代码实现:

---

```

    context_out, _ = self.attention(context, target_mean)
    context_out = self.conv(context_out)
    context_out = context_out.squeeze(dim=1)
[0076]     target_out, _ = self.attention(target,
    context_mean)
    target_out = self.conv(target_out)
    target_out = target_out.squeeze(dim=1)

```

---

[0077] 可选的,步骤S120可以通过如下代码实现:

---

```

    self.attention = Attention(opt.bert_dim,
[0078]     out_dim=opt.hidden_dim, n_head=8, score_function='mlp',
    dropout=opt.dropout)

```

---

[0079] 请参阅图3,示出了本申请实施例中的目标神经网络对句子进行方面级文本情感分类的处理过程示例图。如图3所示,在获取了待处理句子和待处理目标词后,可以将待处理句子和待处理目标词分别输入BERT模型,以获取与句子和目标词分别对应的待处理词向量。进一步的,可以对于与句子和目标词分别对应的待处理词向量进行求平均操作,得到与句子和目标词分别对应的参考词向量。

[0080] 进一步的,可以将与句子和目标词分别对应的词向量、和与句子和目标词分别对应的参考词向量输入目标神经网络的多头注意力网络层,将多头注意力网络层的输出输入到卷积层,对卷积层的输出进行求平均操作,再通过softmax函数对求平均操作后的结果进行分析,进而获取句子的情感极性。

[0081] 步骤S130:基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示。

[0082] 可选的,本实施例中的目标神经网络还可以包括卷积层。作为一种方式,可以将第一多头注意力向量输入卷积层,得到第一隐藏状态;以及将第二多头注意力向量输入卷积层,得到第二隐藏状态;再通过对第一隐藏状态进行求平均操作,可以得到与句子对应的第一隐藏语义表示;以及通过对第二隐藏状态进行求平均操作,可以得到与目标词对应的第二隐藏语义表示。

[0083] 可选的,卷积层的计算公式如下:

[0084]  $\text{Conv}(h) = \text{ELU}(h * W^1 + b^1) * W^2 + b^2$ ,

[0085] 其中, $W^1$ 和 $W^2$ 是权重矩阵, $b^1$ 和 $b^2$ 是偏置项,\*表示卷积操作。

[0086] 作为一种方式,可以将第一多头注意力向量以及第二多头注意力向量输入卷积层,得到隐藏状态,该隐藏状态可以理解为句子和目标词的共同编码注意力信息

$h_s = [h_s^1; h_s^2; \dots; h_s^m]$  和  $h_T = [h_T^i; h_T^{i+1}; \dots; h_T^{i+n-1}]$ , 其中:

[0087]  $h_s = \text{Conv}(s^a)$ 。

[0088] 可选的,通过对 $h_s$ 和 $h_T$ 进行求平均操作,可以得到与句子对应的第一隐藏语义表示

$\bar{h}_S$ 和与目标词对应的第二隐藏语义表示 $\bar{h}_T$ ：

$$[0089] \quad \bar{h}_S = \sum_{i=1}^m \frac{h_S^i}{m},$$

$$[0090] \quad \bar{h}_T = \sum_{i=1}^n \frac{h_T^i}{n}。$$

[0091] 步骤S140：基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量。

[0092] 作为一种方式，可以基于预设规则将第一隐藏语义表示与第二隐藏语义表示进行连接，得到分类特征向量。其中，预设规则可以为：

$$[0093] \quad \gamma = [\bar{h}_S; \bar{h}_T]。$$

[0094] 其中， $\gamma$  表征分类特征向量。

[0095] 步骤S150：基于所述分类特征向量获取所述句子的情感极性类别，以对所述句子进行情感分类。

[0096] 作为一种方式，可以对分类特征向量进行线性特征变换，将线性变换后的结果投影到指定空间，再获取指定空间下句子对应的情感极性类别的预测概率，然后将预测概率最大对应的情感极性类别作为句子的情感极性类别，从而实现对句子进行情感分类。

[0097] 可选的，可以通过公式 $x = W_1 \cdot \gamma + b_1$ 对分类特征向量进行线性特征变换，通过公式

$$P_c = \frac{\exp(x_c)}{\sum_{i \in C} \exp(x_i)}$$

获取指定空间下句子对应的情感极性类别的预测概率。其中，上式中 $W_1$ 为

权重矩阵， $b_1$ 为偏置项， $P_c$ 用于表征情感极性为 $c$ 类的预测概率， $C$ 表示情感极性的类别数。可选的，可以将概率最大的标签对应的情感极性类别作为最终预测的情感极性类别。

[0098] 可选的，情感极性类别可以包括“积极的 (Positive)”、“消极的 (Negative)”或者是“中立的 (Neutral)”。在一些实施方式中，“积极的”情感极性还可进一步包括更细分的情感类别如开心、兴奋、感激等，同理“消极的”还可包括痛苦、仇恨、嫉妒等，具体可以不作限定。

[0099] 本实施例提供的数据处理方法，通过获取待处理词向量，继而获取基于所述待处理词向量确定的多头注意力向量，所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量，所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示，所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示；再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示，以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示；继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量；然后基于所述分类特征向量获取所述句子的情感极性类别，以对所述句子进行情感分类。从而通过上述方式实现了通过获取基于待处理词向量确定的多头注意力向量，以获得与句子和目标词对应的融合了句子的全句语义的向量表示，从而使得分类特征向量的分类结果更加准确，进而可以提升情感分类的准确度。

[0100] 请参阅图4，示出了本申请实施例提供的一种数据处理装置的结构框图，本实施例

提供一种数据处理装置200,可以运行于电子设备,所述装置200可以包括:词向量获取模块210、多头注意力向量获取模块220、隐藏语义表示获取模块230、分类特征向量获取模块240以及情感极性类别获取模块250:

[0101] 词向量获取模块210,用于获取待处理词向量,所述待处理词向量包括与句子和目标词分别对应的词向量、以及与所述句子和所述目标词分别对应的参考词向量。

[0102] 可选的,词向量获取模块210具体可以用于获取待处理句子和待处理目标词,所述待处理句子包括所述待处理目标词;将所述待处理句子和所述待处理目标词输入BERT模型,所述BERT模型为预训练得到,获取所述BERT模型输出的与所述待处理句子和所述待处理目标词分别对应的向量表示,所述向量表示中所述待处理句子和所述待处理目标词对应,所述向量表示融合了全句语义信息;对所述向量表示进行求平均操作,得到待处理词向量。

[0103] 多头注意力向量获取模块220,用于获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示。

[0104] 作为一种方式,多头注意力向量获取模块220可以用于将所述待处理词向量输入目标神经网络的多头注意力网络层,得到所述多头注意力网络层输出的多头注意力向量。其中,可以通过将与所述句子对应的词向量和与所述目标词对应的参考词向量输入多头注意力函数,输出得到第一多头注意力向量;将与所述目标词对应的词向量和与所述句子对应的参考词向量输入多头注意力函数,输出得到第二多头注意力向量。

[0105] 隐藏语义表示获取模块230,用于基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示。

[0106] 可选的,所述目标神经网络还可以包括卷积层,在这种方式下,隐藏语义表示获取模块230具体可以用于将所述第一多头注意力向量输入所述卷积层,得到第一隐藏状态;将所述第二多头注意力向量输入所述卷积层,得到第二隐藏状态;对所述第一隐藏状态进行求平均操作,得到与所述句子对应的第一隐藏语义表示;对所述第二隐藏状态进行求平均操作,得到与所述目标词对应的第二隐藏语义表示。

[0107] 分类特征向量获取模块240,用于基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量。

[0108] 可选的,分类特征向量获取模块240可以用于基于预设规则将所述第一隐藏语义表示与所述第二隐藏语义表示进行连接,得到分类特征向量。

[0109] 情感极性类别获取模块250,用于基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。

[0110] 可选的,情感极性类别获取模块250可以用于对所述分类特征向量进行线性特征变换,将线性变换后的结果投影到指定空间,获取所述指定空间下所述句子对应的情感极性类别的预测概率;将预测概率最大对应的情感极性类别作为所述句子的情感极性类别。

[0111] 本实施例提供的数据处理装置,通过获取待处理词向量,继而获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二

多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;然后基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。从而通过上述方式实现了通过获取基于待处理词向量确定的多头注意力向量,以获得与句子和目标词对应的融合了句子的全句语义的向量表示,从而使得分类特征向量的分类结果更加准确,进而可以提升情感分类的准确度。

[0112] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述装置和模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0113] 在本申请所提供的几个实施例中,所显示或讨论的模块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,装置或模块的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0114] 另外,在本申请各个实施例中的各功能模块可以集成在一个处理模块中,也可以是各个模块单独物理存在,也可以两个或两个以上模块集成在一个模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。

[0115] 请参阅图5,基于上述的数据处理方法及装置,本申请实施例还提供了一种可以执行前述数据处理方法的电子设备100。电子设备100包括存储器102以及相互耦合的一个或多个(图中仅示出一个)处理器104,存储器102以及处理器104之间通信线路连接。存储器102中存储有可以执行前述实施例中内容的程序,而处理器104可以执行存储器102中存储的程序。

[0116] 其中,处理器104可以包括一个或者多个处理核。处理器104利用各种接口和线路连接整个电子设备100内的各个部分,通过运行或执行存储在存储器102内的指令、程序、代码集或指令集,以及调用存储在存储器102内的数据,执行电子设备100的各种功能和处理数据。可选地,处理器104可以采用数字信号处理(Digital Signal Processing,DSP)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)、可编程逻辑阵列(Programmable Logic Array,PLA)中的至少一种硬件形式来实现。处理器104可集成中央处理器(Central Processing Unit,CPU)、图像处理器(Graphics Processing Unit,GPU)和调制解调器等中的一种或几种的组合。其中,CPU主要处理操作系统、用户界面和应用程序等;GPU用于负责显示内容的渲染和绘制;调制解调器用于处理无线通信。可以理解的是,上述调制解调器也可以不集成到处理器104中,单独通过一块通信芯片进行实现。

[0117] 存储器102可以包括随机存储器(Random Access Memory,RAM),也可以包括只读存储器(Read-Only Memory)。存储器102可用于存储指令、程序、代码、代码集或指令集。存储器102可包括存储程序区和存储数据区,其中,存储程序区可存储用于实现操作系统的指令、用于实现至少一个功能的指令(比如触控功能、声音播放功能、图像播放功能等)、用于实现前述各个实施例的指令等。存储数据区还可以存储电子设备100在使用中所创建的数据(比如电话本、音视频数据、聊天记录数据)等。

[0118] 请参考图6,其示出了本申请实施例提供的一种计算机可读存储介质的结构框图。

该计算机可读存储介质300中存储有程序代码,所述程序代码可被处理器调用执行上述方法实施例中所描述的方法。

[0119] 计算机可读存储介质300可以是诸如闪存、EEPROM(电可擦除可编程只读存储器)、EPROM、硬盘或者ROM之类的电子存储器。可选地,计算机可读存储介质300包括非瞬时性计算机可读介质(non-transitory computer-readable storage medium)。计算机可读存储介质300具有执行上述方法中的任何方法步骤的程序代码310的存储空间。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。程序代码310可以例如以适当形式进行压缩。

[0120] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不必针对的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0121] 综上所述,本申请实施例提供了一种数据处理方法、装置、电子设备以及存储介质,通过获取待处理词向量,继而获取基于所述待处理词向量确定的多头注意力向量,所述多头注意力向量包括第一多头注意力向量以及第二多头注意力向量,所述第一多头注意力向量用于表征所述句子对所述目标词的多头注意力向量表示,所述第二多头注意力向量用于表征所述目标词对所述句子的多头注意力向量表示;再基于所述第一多头注意力向量获取与所述句子对应的第一隐藏语义表示,以及基于所述第二多头注意力向量获取与所述目标词对应的第二隐藏语义表示;继而基于所述第一隐藏语义表示以及所述第二隐藏语义表示获取分类特征向量;然后基于所述分类特征向量获取所述句子的情感极性类别,以对所述句子进行情感分类。从而通过上述方式实现了通过获取基于待处理词向量确定的多头注意力向量,以获得与句子和目标词对应的融合了句子的全句语义的向量表示,从而使得分类特征向量的分类结果更加准确,进而可以提升情感分类的准确度。

[0122] 最后应说明的是:以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不驱使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

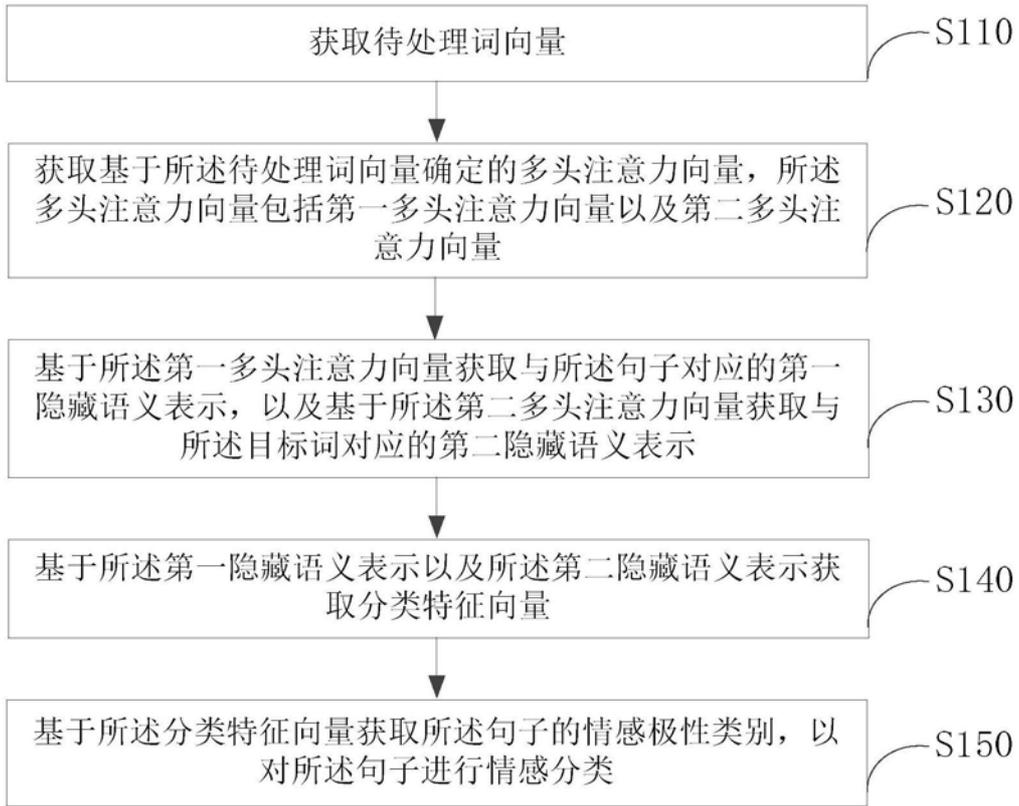


图1

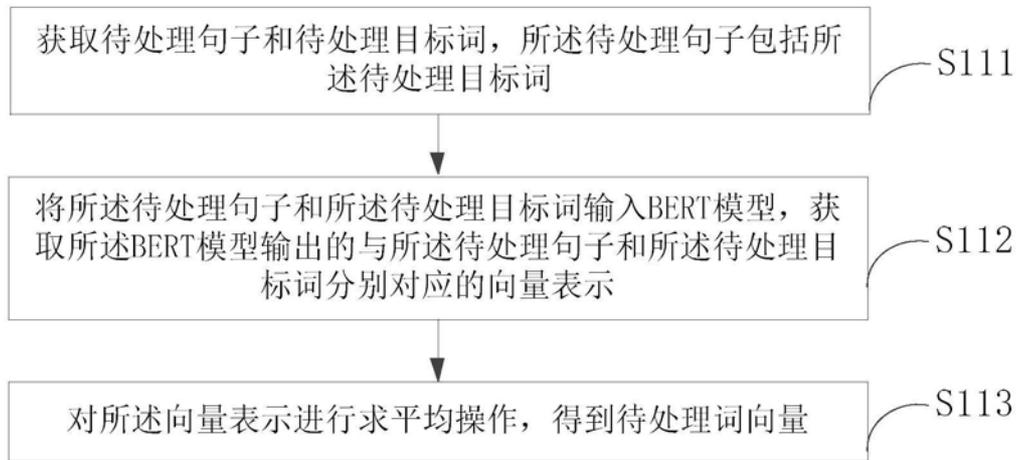


图2

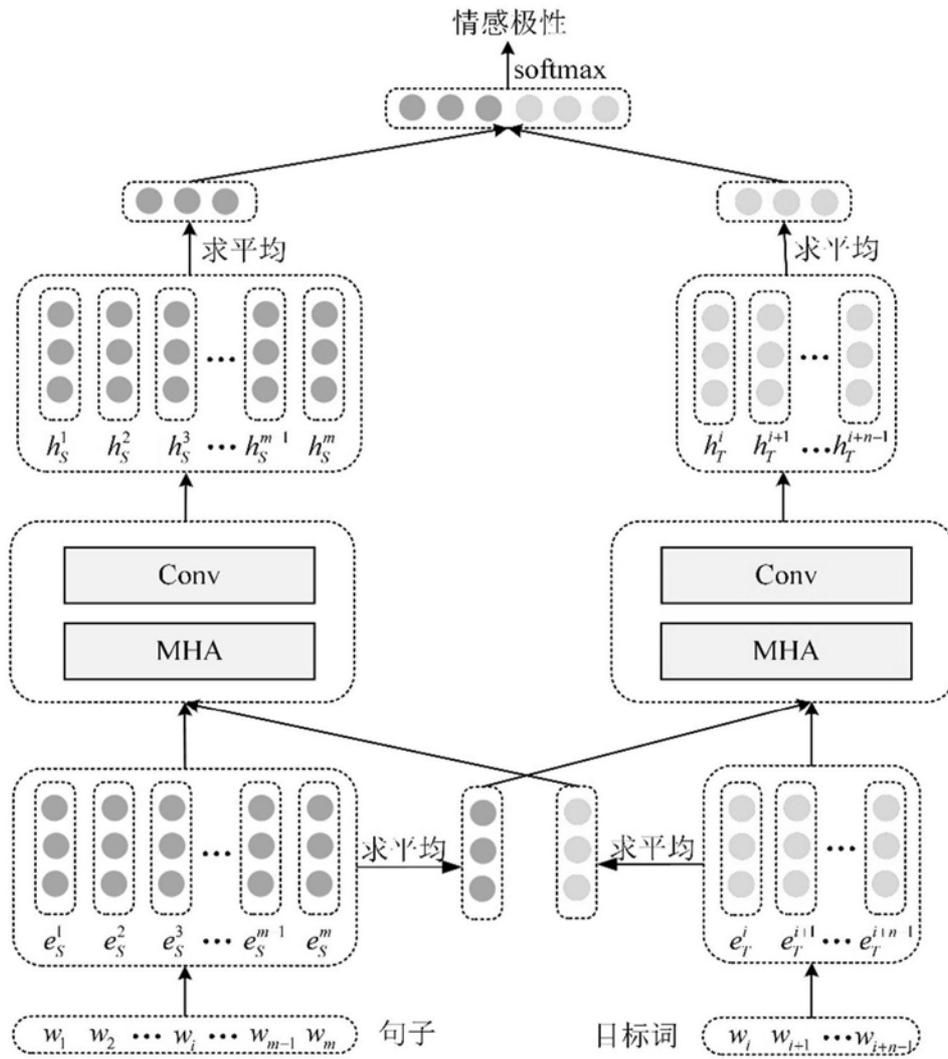


图3



图4

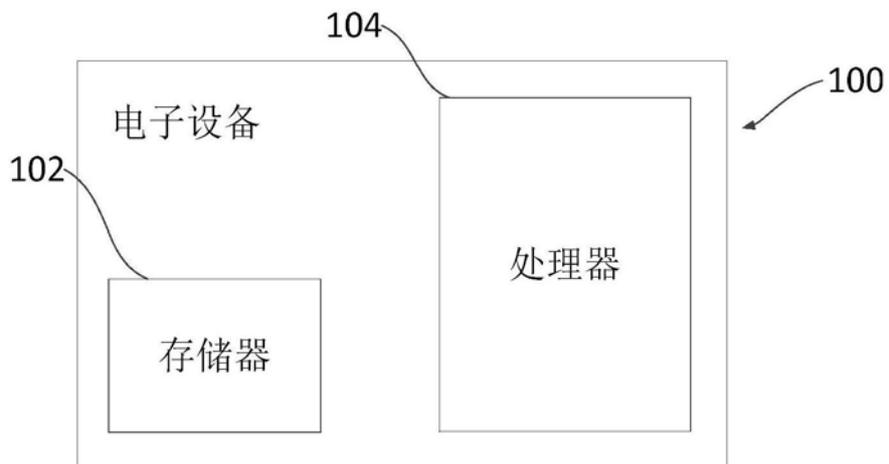


图5



图6