

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局



(10) 国际公布号  
**WO 2021/027533 A1**

(43) 国际公布日  
2021年2月18日 (18.02.2021)

- (51) 国际专利分类号:  
**G06F 17/00** (2019.01)
- (21) 国际申请号: PCT/CN2020/104679
- (22) 国际申请日: 2020年7月25日 (25.07.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201910744603.5 2019年8月13日 (13.08.2019) CN
- (71) 申请人: 平安国际智慧城市科技股份有限公司 (PINGAN INTERNATIONAL SMART CITY TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省深圳市前海深港合作区妈湾兴海大道3048号前海自贸大厦1-34层, Guangdong 518000 (CN)。

自贸大厦1-34层, Guangdong 518000 (CN)。张然 (ZHANG, Ran); 中国广东省深圳市前海深港合作区妈湾兴海大道3048号前海自贸大厦1-34层, Guangdong 518000 (CN)。

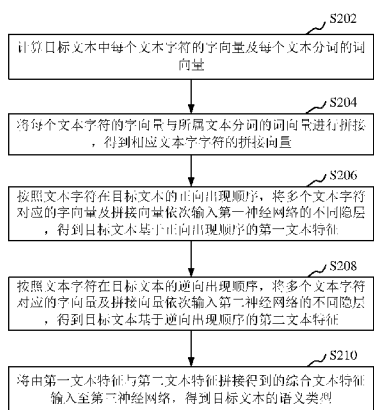
(74) 代理人: 广州三环专利商标代理有限公司 (SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (CN)。

(72) 发明人: 卢清明 (LU, Qingming); 中国广东省深圳市前海深港合作区妈湾兴海大道3048号前海

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: TEXT SEMANTIC RECOGNITION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 文本语义识别方法、装置、计算机设备和存储介质



S202 Calculate a character vector of each text character in a target text and a word vector of each text segmented word

S204 Splice the character vector of each text character with the word vector of the pertaining text segmented words to obtain a corresponding splicing vector of the text characters

S206 According to the forward order of occurrence of the text characters in the target text, sequentially input the character vectors of the multiple text characters and the splicing vectors into different hidden layers of a first neural network to obtain a first text feature of the target text based on the forward order of occurrence

S208 According to the reverse order of occurrence of the text characters in the target text, sequentially input the character vectors corresponding to the multiple text characters and the splicing vectors into different hidden layers of a second neural network to obtain a second text feature of the target text based on the reverse order of occurrence

S210 Input a comprehensive text feature obtained by splicing the first text feature and the second text feature into a third neural network to obtain the semantic type of the target text

图 2

(57) Abstract: A text semantic recognition method and apparatus, a computer device, and a storage medium. The method comprises: calculating a character vector of text characters in a target text and a word vector of each text segmented word; splicing the character vector of each text character with the word vector of the pertaining text segmented words to obtain a splicing vector of the text characters; according to the forward order of occurrence of the text characters in the target text, sequentially inputting the character vectors of the text characters and the splicing vectors into a first neural network to obtain a first text feature; according to the reverse order of occurrence of the text characters in the target text, sequentially inputting the character vectors corresponding to the text characters and the splicing vectors into a second neural network to obtain a second text feature; and inputting a comprehensive text feature obtained by splicing the first text feature and the second text feature into a third neural network to obtain the semantic type of the target text.



WO 2021/027533 A1

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,  
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

---

(57) 摘要: 一种文本语义识别方法、装置、计算机设备和存储介质。方法包括: 计算目标文本中本字符的字向量及每个文本分词的词向量; 将每个文本字符的字向量与所属文本分词的词向量进行拼接得到文本字符的拼接向量; 按照文本字符在目标文本的正向出现顺序, 将文本字符的字向量及拼接向量依次输入第一神经网络得到第一文本特征; 按照文本字符在目标文本的逆向出现顺序, 将文本字符对应的字向量及拼接向量依次输入第二神经网络得到第二文本特征; 将由第一文本特征与第二文本特征拼接得到的综合文本特征输入至第三神经网络得到目标文本的语义类型。

## 文本语义识别方法、装置、计算机设备和存储介质

本申请要求于2019年08月13日提交中国专利局、申请号为201910744603.5，发明名称为“文本语义识别方法、装置、计算机设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

### 技术领域

本申请涉及人工智能技术领域，特别是涉及文本语义识别方法、装置、计算机设备和存储介质。

### 背景技术

随着互联网的发展，文本语义识别技术得到了越来越广泛的应用。尤其是在智能问答领域，为了准确回答用户咨询的问题，通常需要将用户输入的语音转成文本数据，进一步对文本数据进行语义识别，判断文本数据所表达的真实含义，从而准确快速的回答用户所咨询的问题。

在网络平台方面，为了维护网络用语的文明，提升用户的使用体验，通常采用文本语义识别技术对发布在网络上的文本进行语义识别，以此识别出暴力、低俗、敏感话题、商业广告等语义信息的文本。

目前，大部分文本语义分析技术采用关键词匹配方法进行处理，需要预先构建关键词数据库，将待识别的文本与已构建的数据库中的关键词进行匹配以此识别出敏感词。发明人意识到，对于数据库中未记录的关键词则无法准确识别其语义，也就是说关键词的覆盖范围限制了文本语义识别的准确率，从而使得文本语义识别的准确率较低。

### 发明内容

基于此，有必要针对上述技术问题，提供一种文本语义识别方法、装置、计算机设备和存储介质。

一种文本语义识别方法，所述方法包括：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

一种文本语义识别装置，所述装置包括：

向量计算模块，用于计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

向量拼接模块，用于将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

第一文本特征获取模块，用于按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

第二文本特征获取模块，用于按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本

基于逆向出现顺序的第二文本特征；

语义类型获取模块，用于将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

一种计算机设备，包括存储器和处理器，所述处理器、和所述存储器相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器用于执行所述存储器的所述程序指令，其中：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令被处理器执行时，用于实现以下步骤：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

上述文本语义识别方法、装置、计算机设备和存储介质，通过计算获得每个文本字符对应的字向量和所属文本分词的词向量并进行向量拼接，得到与文本字符对应的拼接向量，通过对文本字符进行向量拼接，通过多种特征向量来表征文本，增强了文本语言表示的特征维度。进一步，通过将字向量和拼接向量按照正向顺序和逆向顺序输入至不同神经网络的不同隐层，可以更充分的获取文本字符的相关信息，挖掘文本字符之间的上下文语义，使得经第一神经网络输出的第一文本特征和第二文本特征进行拼接得到综合特征，可以更充分的表达目标文本的语义特征，提高了文本语义识别的准确率。

## 附图说明

图1为一个实施例中文本语义识别方法的应用场景图；

图2为一个实施例中文本语义识别方法的流程示意图；

图3为一个实施例中预设文件生成的流程示意图；

图4为一个实施例中文本语义识别装置的结构框图；

图5为另一个实施例中文本语义识别装置的结构框图；

图6为一个实施例中计算机设备的内部结构图。

## 具体实施方式

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

本申请提供的文本语义识别方法，可以应用于如图 1 所示的应用环境中。该文本语义识别方法应用于文本语义系统中。该文本语义系统包括终端 102 和服务器 104。其中，终端 102 与服务器 104 通过网络进行通信。文本语义识别方法可以在终端 102 或服务器 104 完成，终端 102 可以采集待识别的目标文本并在终端 102 上采用上述文本语义识别方法进行语义类型的识别。或者终端 102 可以获取待识别的目标文本后，通过网络连接将目标文本传输至服务器 104，服务器 104 采用上述文本语义识别方法对目标文本进行语义类型的识别。其中，终端 102 可以但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备，服务器 104 可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在一个实施例中，如图 2 所示，提供了一种文本语义识别方法，以该方法应用于图 1 中的服务器为例进行说明，包括以下步骤：

步骤 S202，计算目标文本中每个文本字符的字向量及每个文本分词的词向量。

其中，文本字符是由目标文本切分得到的多个独立字符。文本字符具体可以是字母、数字、字或符号等。文本分词是指将目标文本切分成一个一个单独的词，即将连续的字符序列按照一定的规范重新组合成词序列的过程。文本分词可以采用基于字符串匹配的分词方法、基于语义的分词方法和基于统计的分词方法进行分词。字向量和词向量用于表征目标文本的多维度表示形式。

具体地，服务器根据获取到的目标文本，确定目标文本中所包含的每个文本字符以及每个文本字符所属的文本分词，通过预先训练好的字向量库或词向量库，匹配得到目标文本中每个文本字符对应的字向量和文本分词对应的词向量。服务器也可以通过预设的向量编码规则对获得的文本字符和文本分词进行编码，得到对应的字向量和词向量。

在其中一个实施例中，获取目标文本的具体步骤包括：终端获取目标文本，其中目标文本可以有多个，该目标文本可以是经语音识别得到的识别文本，也可以是用户直接在终端输入的文本。终端将获取的目标文本传输至服务器。目标文本也可以是从网络平台上获取的，通过爬虫技术从网络上获取相关目标文本。

在其中一个实施例中，确定目标文本中所包含的每个文本字符以及每个文本字符所属的文本分词的步骤包括：服务器根据接收到的目标文本按字符进行分字处理，得到目标文本所包含的文本字符；将得到的各个文本字符按照文本字符在目标文本中出现的先后顺序进行排列，得到目标文本的字符序列，从该字符序列中删除属于停用词表的文本字符，获得经过预处理后的字符序列。其中停用词是指在自然语言处理任务中需要被过滤的不具有处理价值的词或字；停用词包括英文字符、数字、数学字符、标点符号及使用频率较高的单汉字等。

服务器对字符序列中的每个字符进行检测，将相同的字符进行字符标识，以区分相同字符所对应的不同词；利用预先构建的分词词库，对具有字符标识的字符序列进行分词处理，得到具有字符标识的词序列；基于经过预处理后的字符序列，服务器从词序列中确定每个字符所属的文本分词。

在其中一个实施例中，构建分词词库时可以通过《新华词典》或者其他类似的出版书籍为基础而建立的词库，也可以根据智能客服场景构建分词词库。已构建的分词词库可存储在服务器的数据库内或发送至云端。

在其中一个实施例中，目标文本也可以通过服务器获取，比如，服务器可以从网页中

获取所需文本数据作为目标文本，进一步确定目标文本的各个文本字符和每个文本字符所属的文本分词。

例如，获取的目标文本为“深圳市的市政府在市民中心。”，首先服务器对目标文本进行分字处理，得到字符序列“深/圳/市/的/市/政/府/在/市/民/中/心/。”，删除字符序列中属于停用词表的字符，得到经过预处理后的字符序列“深/圳/市/市/政/府/市/民/中/心/”；进一步，将相同的字符进行字符标识，即得到具有字符标识的字符序列“深/圳/市 01/市 02/政/府/市 03/民/中/心/”，对该字符序列进行分词，得到词序列“深圳市 01/市 02 政府/市 03 民中心”，虽然文本字符为“市”对应有三个词，但是可以依据字符标识区分文本字符所属的文本分词。

步骤 S204，将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字符的拼接向量。

其中，拼接向量是将多个文本向量按预设规则拼接而形成的一个向量，拼接向量表征了多个向量的表示维度。

具体地，基于获取的目标文本的字向量和词向量，服务器将每个文本字符对应的字向量和该文本字符所属的词向量进行拼接，得到与该文本字符对应的拼接向量，以此获得目标文本所包含的所有文本字符的拼接向量，其中字向量和词向量进行拼接的顺序不做要求。

在其中一个实施例，服务器将每个文本字符对应的字向量以及与文本字符所属文本分词的词向量进行相加或相乘，得到相应文本字符的拼接向量。

步骤 S206，按照文本字符在目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到目标文本基于正向出现顺序的第一文本特征。

其中，第一神经网络主要用于将按照文本字符在目标文本的正向出现顺序输入的目标文本所包含的特征，生成携带有基于正向出现顺序的目标文本的上下文语义的特征；第一神经网络包括多个隐层，每个隐层可以具有相同或不同的神经元节点。第一神经网络是递归神经网络，可以是长短期记忆网络 LSTM 和循环神经网络 RNN。

具体地，目标文本包含多个文本字符，服务器计算每个文本字符对应的字向量和拼接向量，将计算得到的字向量和拼接向量按照文本字符在目标文本的正向出现顺序进行排序，进一步，将按照正向出现顺序排序的字向量和拼接向量依次输入至第一神经网络的不同隐层中进行特征提取，获取不同文本字符的相互信息，得到基于正向出现顺序的第一文本特征。

在其中一个实施例，可以将多个文本字符对应的字向量、词向量和拼接向量依次输入至第一神经网络的不同隐层，得到目标文本基于正向出现顺序的第一文本特征。

步骤 S208，按照文本字符在目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到目标文本基于逆向出现顺序的第二文本特征。

其中，第二神经网络主要用于将按照文本字符在目标文本的逆向出现顺序输入的目标文本所包含的特征，生成携带有基于逆向出现顺序的目标文本的上下文语义的特征；第二神经网络包括多个隐层，每个隐层可以具有相同或不同的神经元节点。第二神经网络是递归神经网络，可以是长短期记忆网络 LSTM 和循环神经网络 RNN。

具体地，基于得到的目标文本中所有文本字符各自对应的字向量和拼接向量，根据文本字符在目标文本中出现的逆向顺序，将文本字符对应的字向量和拼接向量依次输入至第二神经网络的不同隐层中，通过第二神经网络对输入的字向量和拼接向量进行特征提取，得到基于逆向出现顺序的第二文本特征。

在其中一个实施例，可以预先设定第一神经网络或第二神经网络可以输入的最大文

本字符的数量，若当前输入的目标文本的文本字符数量小于最大文本字符的数量，则在目标文本所形成的字向量矩阵中用 0 向量补齐，将补齐后的字向量矩阵作为第一神经网络或第二神经网络的输入。

步骤 S210，将由第一文本特征与第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到目标文本的语义类型。

其中，综合文本特征是将经第一神经网络的输出和第二神经网络输出，按预设规则拼接而形成的一个文本特征。第三神经网络主要用于根据输入的目标文本对应的综合文本特征按语义类型进行分类，以此得到目标文本的语义类型。语义类型是指根据目标文本的语义关系确定目标文本所属的类型。

具体地，基于得到的目标文本的第一文本特征和第二文本特征，服务器将第一文本特征和第二文本特征进行拼接得到目标文本的综合文本特征；进一步，将目标文本的综合文本特征传输至第三神经网络，通过第三神经网络对综合文本特征按语义类进行分类，得到目标文本的语义类型，充分考虑了对文本上下文和隐含词汇的语义理解。比如，在识别目标文本的脏话和礼貌用语时，可以相应将语义类型设置为两类，即类别 1：文本是脏话，类别 0：文本是礼貌用语。

在上述实施例中，通过计算获得每个文本字符对应的字向量和所属文本分词的词向量并进行向量拼接，得到与文本字符对应的拼接向量，通过对文本字符进行向量拼接，通过多种特征向量来表征文本，增强了文本语言表示的特征维度。进一步，通过将字向量和拼接向量按照正向顺序和逆向顺序输入至不同神经网络的不同隐层，可以更充分的获取文本字符的相关信息，挖掘文本字符之间的上下文语义，使得经第一神经网络输出的第一文本特征和第二文本特征进行拼接得到综合特征，可以更充分的表达目标文本的语义特征，提高了文本语义识别的准确率。

在一个实施例中，如图 3 所示，上述方法还包括预设文件生成的步骤：

步骤 S302，获取样本文本。

步骤 S304，基于预训练的第一神经网络提取样本文本的字向量及词向量。

步骤 S306，对字向量和词向量分别进行字符编号。

步骤 S308，将字向量、词向量以及分别对应的字符编号写入到预设文件。

计算文本字符的字向量和文本分词对应的词向量包括：对每个文本字符以及文本分词进行字符编号；基于字符编号，在预设文件中读取得到每个文本字符对应的字向量以及每个文本分词对应的词向量。

其中，预设文件是预先构建的具有索引的文本，包括字向量及其索引，词向量及其索引。

具体地，在计算文本字符的字向量和文本分词对应的词向量之前，需构建包含字向量和词向量的索引查询的预设文件。服务器从终端或网页上获取到样本文本以及对应的已知语义类型，基于预先训练好的第一神经网络提取样本文本的字向量和词向量，并对提取的字向量和词向量分别进行字符编号，得到字向量与数字的映射关系，以及词向量和数字的映射关系。服务器将字向量、词向量以及分别对应的字符编号写到预设文件中，形成具有字符编号索引的字向量和词向量。

基于得到的目标文本所包含的各个文本字符以及与文本字符所属的文本分词，服务器对每个文本字符和文本分词分别进行字符编号，获得文本字符与字符编号间的映射关系，以及文本分词与字符编号间的映射关系。根据每个文本字符的字符编号从预设文件中查询得到相应的文本字符的字向量，以及根据每个文本分词的字符编号从预设文件中查询得到相应文本分词的词向量。

在其中一个实施例中，字符编号可以包括编号类型。根据编号类型对字向量和词向量

分别进行字符编号，字向量和词向量的编号类型可以相同也可以不同。比如，对字向量按照自然数进行字符编号；对于词向量可以按照自然数进行编号，也可以按照英文字母进行编号。

比如，目标文本为“深圳市”，假如文本字符“深”的字符编号为01，从预设文件中查询得到字符编号为01所对应的字向量是(1,1,2,2)。

在本实施例中，通过预先构建的包含字向量和词向量的预设文件，在计算目标文本的字向量和词向量时，根据文本字符与文本分词各自对应的字符编号，从预设文件中查询得到各自对应的字向量和词向量，可以准确快速的得到文本字符的字向量以及文本分词的词向量，从而提高了获取目标文本的语义类型的速率和准确率。

在一个实施例中，按照文本字符在目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到目标文本基于正向出现顺序的第一文本特征包括：按照文本字符在目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层；将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层；将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到目标文本基于正向出现顺序的第一文本特征。

具体地，根据文本字符在目标文本中的正向出现顺序，服务器将位于第一顺序的文本字符对应的字向量和拼接向量输入至第一神经网络的第一隐层中，通过第一隐层对输入的字向量和拼接向量进行投影得到该文本字符对应的字符特征。其中，第一顺序文本字符是指目标文本中的所有文本字符按预设的出现顺序进行排序后，位于第一位的文本字符即为第一顺序文本字符。最后顺序文本字符是指目标文本中的所有文本字符按预设的出现顺序进行排序后，位于最后一位的文本字符。

服务器将第一隐层作为当前隐层，按照文本字符在目标文本的正向出现顺序，得到当前隐层对应的字向量和拼接向量。预先设置当前隐层内各个神经元节点的权重，其中，每个神经元节点的权重可以相同也可以不同。进而，服务器根据预先设置的各个神经元节点的权重对当前隐层输入的特征进行非线性映射，得到当前隐层输出的字符特征。其中，非线性映射可以采用sigmoid(S型)函数、tanh(双曲正切)函数、relu(修正线性单元)函数等激活函数。

服务器将当前隐层输出的字符特征与下一顺序文本字符对应的字向量和拼接向量输入至第一神经网络的下一隐层，将下一隐层作为当前隐层，循环执行以下步骤：按照文本字符在目标文本的正向出现顺序，得到当前隐层对应的字向量和拼接向量。预先设置当前隐层内各个神经元节点的权重。服务器根据当前隐层对应的字向量和拼接向量，当前隐层对应的各个神经元节点的权重和上一隐层输出的字符特征，采用非线性映射得到当前隐层输出的字符特征。即服务器将上一隐层输出的字符特征、当前隐层对应的字向量和拼接向量输入至当前隐层，根据预先设置的各个神经元节点的权重对输入的特征进行非线性映射，以此得到下一隐层的字符特征。将下一隐层作为当前隐层，重复进行迭代，执行上述步骤，直至将最后顺序文本字符对应的字向量和拼接向量输入至当前隐层，得到当前隐层输出的字符特征，将该输出的字符特征作为目标文本基于正向出现顺序的第一文本特征。

在其中一个实施例中，当前隐层包括第一子隐层和第二子隐层；将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层包括：将字向量和上一隐层的输出作为第一子隐层的输入，第一子隐层用于根据第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；将第一子字符特征和拼接向量作为第二子隐层的输入，第二子隐层用于根据第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

具体地，当前隐层包括第一子隐层和第二子隐层；服务器将出现顺序位于第一位的文



本字符对应的字向量输入至第一子隐层，并通过第一子隐层内预先设置的各个神经元节点的权重对该字向量进行投影，得到第一子隐层输出的第一字符特征；进一步，服务器将第一字符特征与出现顺序位于第一位的文本字符对应的拼接向量作为第二子隐层的输入，根据预先设置的第二子隐层中各个神经元节点的权重对输入的特征进行非线性映射，以此得到第二子隐层输出的字符特征，作为第一隐层的输出。

在其中一个实施例中，当前隐层包括第一子隐层和第二子隐层：将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层包括：服务器将拼接向量和上一隐层的输出作为第一子隐层的输入，第一子隐层用于根据第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；将第一子字符特征和字向量作为第二子隐层的输入，第二子隐层用于根据第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

在其中一个实施例中，按照文本字符在目标文本的正向出现顺序，服务器将多个文本字符对应的字向量、词向量和拼接向量依次输入至第一神经网络的不同隐层，得到目标文本基于正向出现顺序的第一文本特征。

具体地，按照文本字符在目标文本的正向出现顺序，服务器将当前顺序文本字符对应的字向量、词向量及拼接向量输入第一神经网络的当前隐层，根据当前隐层对应的字向量、词向量和拼接向量，当前隐层对应的各个神经元节点的权重和上一隐层输出的字符特征，采用非线性映射得到当前隐层输出的字符特征；服务器将当前隐层输出的字符特征与下一顺序文本字符对应的字向量、词向量及拼接向量输入第一神经网络的下一隐层，将下一隐层作为当前隐层，重复进入将当前隐层输出的字符特征与下一顺序文本字符对应的字向量、词向量及拼接向量输入第一神经网络的下一隐层的步骤，直至最后顺序文本字符，得到目标文本基于正向出现顺序的第一文本特征。

在其中一个实施例中，第一隐层包括第一子隐层、第二子隐层和第三子隐层：将当前顺序文本字符对应的字向量、词向量及拼接向量输入第一神经网络的当前隐层包括：将字向量和上一隐层的输出作为第一子隐层的输入，第一子隐层用于根据第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；将第一子字符特征和词向量作为第二子隐层的输入，第二子隐层用于根据第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征；将第二子字符特征和拼接向量作为第三子隐层的输入，第三子隐层用于根据第三子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第三子字符特征，作为第一隐层的输出。其中，文本字符对应的字向量、词向量以及拼接向量进行投影的顺序不做规定，可以任意设置。比如，可以将词向量进行投影得到第一子隐层对应的第一子字符特征。

在其中一个实施例中，服务器根据当前隐层对应的字向量和拼接向量、当前隐层对应的各个神经元节点的权重和上一隐层输出的字符特征，采用非线性映射得到下一隐层的字符特征具体可以通过下述例子说明。

例如，假设当前隐层对应的各个神经元节点记为 $W_f$ ，当前隐层对应的字向量为 $x_i$ ，当前隐层对应的拼接向量 $y_i$ ，上一隐层输出的字符特征为 $h_{i-1}$ ，非线性函数为 $\tanh$ ，则下一隐层输出的字符特征 $f_i$ 可以用以下公式计算得到：

$$f_i = \tanh(W_c [h_{i-1}, x_i, y_i] + b_f) ; \text{其中, } b_f \text{ 为当前隐层的偏置。}$$

在其中一个实施例中，第一神经网络还包括随机失活层，方法还包括：将第一文本特征作为随机失活层的输入，随机失活层用于将第一文本特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征向量，作为第一神经网络的输出。

其中，随机失活层(dropout)主要用于对输入的第一文本特征进行稀疏处理，将第

一文本特征的部分元素进行归零处理，防止神经网络的过拟合，同时也降低神经网络的计算量。

具体地，服务器将第一文本特征输入至随机失活层，随机失活层根据设定的稀疏概率对第一文本特征进行稀疏处理，将第一文本特征中每个数据按照稀疏概率进行投影，以此得到稀疏特征向量，其中稀疏概率是指数据进行投影后所出现的概率。比如，第一文本特征为一维序列 $[1,2,3,4]^T$ ，设定稀疏概率是0.5，对应的一维序列中的每个数字投影后出现的概率都是0.5，即经过随机失活层所输出的结果可以是 $[0,2,0,4]^T$ ，也可以是 $[0,0,0,4]^T$ 。

在一个实施例中，按照文本字符在目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到目标文本基于逆向出现顺序的第二文本特征包括：按照文本字符在所述目标文本的逆向出现顺序，将所述第一文本特征、第一顺序文本字符对应的字向量和拼接向量输入至第二神经网络的第一隐层，得到所述第一隐层输出的字符特征；将第二神经网络的第二隐层作为当前隐层，将第二顺序文本字符作为当前顺序文本字符；将所述前一隐层输出的字符特征、当前顺序文本字符对应的字向量及拼接向量输入第二神经网络的当前隐层；将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于逆向出现顺序的第二文本特征。

具体地，将第一神经网络和第二神经网络串联，将第一神经网络的输出作为第二神经网络的输入。按照文本字符在目标文本的逆向出现顺序，获得第一顺序文本字符对应的字向量和拼接向量；服务器将获得的第一神经网络输出的第一文本特征、以及该字向量和拼接向量输入至第二神经网络的第一隐层，得到第一隐层输出的字符特征；将第二神经网络的第二隐层作为当前隐层，将第二顺序文本字符作为当前顺序文本字符；进一步，将前一隐层输出的字符特征、当前顺序文本字符对应的字向量和拼接向量输入至第二神经网络的当前隐层，通过当前隐层内设置的各个神经元的权重对输入的特征进行非线性映射，得到当前隐层输出的字符特征；服务器将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第二神经网络的下一隐层，将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到目标文本基于逆向出现顺序的第二文本特征。

在其中一个实施例中，第二神经网络的当前隐层包括第一子隐层和第二子隐层；将所述前一隐层输出的字符特征、当前顺序文本字符对应的字向量及拼接向量输入第二神经网络的当前隐层包括：将字向量和前一隐层输出的字符特征作为第一子隐层的输入，第一子隐层用于根据第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；将第一子字符特征和拼接向量作为第二子隐层的输入，第二子隐层用于根据第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

在其中一个实施例中，将第二文本特征输入至第三神经网络，得到目标文本的语义类型包括：将第二文本特征作为注意力机制层的输入，注意力机制层用于将第二文本特征中每个数据进行加权处理得到加权特征；将加权特征作为随机失活层的输入，随机失活层用于将加权特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征；将稀疏特征作为全连接层的输入，全连接层用于对稀疏特征进行分类运算得到每个语义类型对应的预测概率；选取预测概率最大的语义类型作为目标文本的语义类型。

在其中一个实施例中，服务器按照文本字符在目标文本的逆向出现顺序，将多个文本字符对应的字向量、词向量和拼接向量依次输入至第二神经网络的不同隐层，得到目标文本基于逆向出现顺序的第二文本特征。

在其中一个实施例中，第二神经网络还包括随机失活层，方法还包括：服务器将第二文本特征作为随机失活层的输入，随机失活层用于将第二文本特征中每个数据按照预设的

稀疏概率进行投影得到稀疏特征向量，作为第二神经网络的输出。

下面，以一个具体的实施例来描述第一神经网络和第二神经网络对目标文本的处理过程。例如，对于目标文本为“深圳市”，文本字符“深”对应的字向量为“(1,1,2,2)”、拼接向量为“(1,1,1,1)”，相应的文本字符“圳”对应“(1,2,3,4)”、拼接向量为“(2,1,1,1)”，文本字符“市”对应“(0,0,2,5)”、拼接向量为“(3,1,1,1)”；

按照文本字符在目标文本的正向出现顺序，形成的目标文本的字向量为  $\begin{pmatrix} 1,1,2,2 \\ 1,2,3,4 \\ 0,0,2,5 \end{pmatrix}$ ，拼接

向量为  $\begin{pmatrix} 1,1,1,1 \\ 2,1,1,1 \\ 3,1,1,1 \end{pmatrix}$ 。

将第一顺序文本字符对应的字向量和拼接向量输入至第一隐层中，并将第一隐层作为当前隐层，当前隐层包括第一子隐层和第二子隐层；即将字向量(1,1,2,2)输入至当前隐层的第一子隐层中，得到第一子隐层输出的第一子字符特征；进一步，将第一子字符特征和拼接向量(1,1,1,1)输入至第二子隐层，根据预先设置的第二子隐层中各个神经元节点的权重对输入的特征进行非线性映射，以此得到第二子隐层输出的字符特征，作为当前隐层的输出。

进一步，将文本字符按正向出现顺序推移得到相应的字向量(1,2,3,4)和拼接向量(2,1,1,1)，并将第二隐层作为当前隐层。将前一隐层的输出和字向量(1,2,3,4)输入至当前隐层的第一子隐层中，得到当前隐层的第一子隐层的输出；将当前隐层的第一子隐层的输出和拼接向量(2,1,1,1)输入至当前隐层的第二子隐层，根据预先设置的当前隐层的第二子隐层中各个神经元节点的权重对输入的特征进行非线性映射，以此得到当前隐层的第二子隐层输出的字符特征，作为当前隐层的输出。直至将前一隐层的输出、最后顺序文本字符对应的字向量(0,0,2,5)和拼接向量(3,1,1,1)输入至当前隐层，得到当前隐层的输出，将该输出结果作为第一神经网络的输出的第一文本特征。

在其中一个实施例中，将第一神经网络和第二神经网络串联，将第二神经网络输出的第二文本特征输入至第三神经网络；其中，第一神经网络的输出作为第二神经网络的输入。

基于文本字符在目标文本的逆向出现顺序，形成的目标文本的字向量为  $\begin{pmatrix} 0,0,2,5 \\ 1,2,3,4 \\ 1,1,2,2 \end{pmatrix}$ ，

拼接向量为  $\begin{pmatrix} 3,1,1,1 \\ 2,1,1,1 \\ 1,1,1,1 \end{pmatrix}$ 。

按照文本字符的逆向出现顺序，将第一文本特征、第一顺序文本字符对应的字向量和拼接向量输入至第二神经网络的第一隐层中，得到所述第一隐层输出的字符特征，并将第

二神经网络的第一隐层作为当前隐层，将第二顺序文本字符作为当前顺序文本字符；即将第一文本特征、字向量 $(0,0,2,5)$ 输入至第一子隐层中，得到当前隐层的第一子隐层的输出。将当前隐层的第一子隐层的输出和拼接向量 $(3,1,1,1)$ 输入至当前隐层的第二子隐层，根据预先设置的当前隐层的第二子隐层中各个神经元节点的权重对输入的特征进行非线性映射，以此得到当前隐层的第二子隐层输出的字符特征，作为当前隐层的输出。

将文本字符按逆向出现顺序推移得到相应的字向量 $(1,2,3,4)$ 和拼接向量 $(2,1,1,1)$ ，并将第二隐层作为当前隐层。将前一隐层的输出和字向量 $(1,2,3,4)$ 输入至当前隐层的第一子隐层中，得到当前隐层的第一子隐层的输出；将当前隐层的第一子隐层的输出和拼接向量 $(2,1,1,1)$ 输入至当前隐层的第二子隐层，根据预先设置的当前隐层的第二子隐层中各个神经元节点的权重对输入的特征进行非线性映射，以此得到当前隐层的第二子隐层输出的字符特征，作为当前隐层的输出。直至将前一隐层的输出、最后顺序文本字符对应的字向量 $(1,1,2,2)$ 和拼接向量 $(1,1,1,1)$ 输入至当前隐层，得到当前隐层的输出，将该输出结果作为第一神经网络输出的第二文本特征。将第二文本特征输入至第三神经网络，得到目标文本的语义类型。通过第一神经网络和第二神经网络的串行连接，将第一神经网络的输出作为第二神经网络的输入，即将基于正向顺序的第一文本特征、以及基于逆向出现顺序的字向量和拼接向量一起输入至第二神经网络，可以更充分的挖掘文本字符之间的相互信息，尤其是在文本字符间隔较远的情况下也能很好的获得字符间的上下文信息。

在其中一个实施例中，第一神经网络和第二神经网络并联，将第一神经网络输出的第一文本特征和第二神经网络输出的第二文本特征进行拼接输入至第三神经网络。通过第一神经网络和第二神经网络对目标文本进行并行处理，可以提高数据处理的速率。

在本实施例中，通过将文本字符对应的字向量、词向量以及拼接向量等多种特征输入至第一神经网络和第二神经网络中，通过第一神经网络进和第二神经网络内各个隐层对输入的特征进行循环计算，以此得到表征目标文本语义的第一文本特征和第二文本特征，通过神经网络的循环计算，可以更好的抓取文本字符之间的相关性，尤其对于字符间隔较远的文本字符。也就是说当文本字符间隔较大时，也能很好的获取预测位置的相关信息。相关信息并不会随着循环次数递增，呈现衰减的趋势。

在一个实施例中，第三神经网络层包括注意力机制层和全连接层；将第一文本特征和第二文本特征进行拼接输入至第三神经网络，得到目标文本的语义类型包括：将第一文本特征和第二文本特征进行拼接，得到目标文本的综合文本特征；将综合文本特征作为注意力机制层的输入，注意力机制层用于将综合文本特征中每个数据进行加权处理得到加权特征；将加权特征作为随机失活层的输入，随机失活层用于将加权特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征；将稀疏特征作为全连接层的输入，全连接层用于对稀疏特征进行分类运算得到每个语义类型对应的预测概率；选取预测概率最大的语义类型作为目标文本的语义类型。

具体地，服务器将目标文本的第一文本特征和第二文本特征进行拼接，以此得到目标文本的综合文本特征。进一步，服务器将得到的综合文本特征输入至注意力机制层，注意力机制层根据预先训练好的系数权重参数计算综合文本特征中各个数据得到系数序列；采用非线性激活函数对系数序列进行激活，得到激活后的系数序列；通过逻辑回归 (softmax) 函数对激活后的系数序列进行归一化处理，得到与综合文本特征中各个数据对应的系数概率；其中，系数概率的范围位于 $[0, 1]$ 之间。将得到的系数概率与综合文本特征各自对应的数据进行相乘，得到加权处理后的加权特征，作为注意力机制层的输出。

进一步，服务器将注意力机制层输出的加权特征输入至随机失活层，随机失活层根据设定的稀疏概率对加权特征进行稀疏处理，将加权特征中每个数据按照稀疏概率进行投影，以此得到稀疏特征，其中稀疏概率是指数据进行投影后所出现的概率。

进一步，服务器将稀疏特征输入至全连接层，通过全连接层对稀疏特征进行分类运算，根据训练好的全连接层的权重参数计算每个语义类型所对应的预测概率，其中全连接层输出的每个预测概率对应一个语义类型。服务器选取预测概率最大的语义类型作为目标文本的语义类型。

在其中一个实施例中，第三神经网络层还包括逻辑回归层 (softmax 层)，具体包括：将每个语义类型对应的预测概率作为 softmax 层的输入，其中 softmax 层用于对各个预测概率进行归一化处理得到每个语义类型对应的概率，选取概率最大的语义类型作为目标文本的语义类型。

例如，经过全连接层输出预测概率为  $\begin{bmatrix} a \\ b \end{bmatrix}$ ， $a$  对应的语义类型是 1， $b$  对应的语义类型是 0；采用 softmax 函数进行归一化后，得到每个语义类型经过归一化后输出的概率为  $\begin{bmatrix} e^a / (e^a + e^b) \\ e^b / (e^a + e^b) \end{bmatrix}$ ，选取最大概率所对应的语义类型作为目标文本的语义类型。

在其中一个实施例中，神经网络包括第一神经网络、第二神经网络和第三神经网络，神经网络模型的训练过程包括：获取样本文本以及已知标签，确定样本文本所包含的样本文本字符以及每个样本文本字符所属的样本文本分词，计算样本文本字符对应的样本字向量和样本文本分词对应的样本词向量，并将每个样本文本字符对应的样本字向量与所属的样本文本分词的样本词向量拼接，得到相应的样本文本字符的样本拼接向量；按照样本文本字符在样本文本的正向出现顺序，将多个样本文本字符对应的样本字向量和样本拼接向量依次输入至待训练的第一神经网络得到样本第一文本特征；按照样本文本字符在样本文本的逆向出现顺序，将多个样本文本字符对应的样本字向量和样本拼接向量依次输入至待训练的第二神经网络得到样本第二文本特征；

将样本第一文本特征与样本第二文本特征进行拼接得到的样本综合文本特征输入至待训练的第三神经网络，得到样本文本的预测语义类型；根据预测语义类型与已知标签计算损失值，将损失值通过反向梯度传播方法传输到神经网络模型各层，获得各层参数的梯度；根据梯度调整神经网络模型中各层的参数，直至所确定的损失值达到训练停止条件。

其中，调整神经网络模型中各层的参数具体包括调整全连接层的权重参数，第一神经网络和第二神经网络中各隐层的权重参数和偏置参数。计算损失值的函数可以是交叉熵损失函数。反向梯度传播方法可以是批量梯度下降方法 (BGD)、小批量梯度下降方法 (MGBD) 以及随机梯度下降方法 (SGD)。

在本实施例中，通过第三神经网络层的注意力机制层对综合文本特征进行加权处理，突出对文本字符之间相互信息较高的特征，并弱化相互信息较低的特征；进一步，通过随机失活层对加权后的特征进行稀疏处理，得到经过稀疏处理后的稀疏特征，并通过全连接层对稀疏特征进行分类运算得到每个语义类型对应的预测概率，选取最大预测概率对应的语义类型作为目标文本的语义类型，利用综合文本特征表征目标文本的语义特征，并通过加权和稀疏处理，增加了文本字符的上下文语义，降低了计算机设备的计算量，同时提高了目标样本的分类准确率。

应该理解的是，虽然图 2-3 的流程图中的各个步骤按照箭头的指示依次显示，但是这

些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2-3 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

在一个实施例中，如图 4 所示，提供了一种文本语义识别装置 400，包括：向量计算模块 402、向量拼接模块 404、第一文本特征获取模块 406、第二文本特征获取模块 408 和语义类型获取模块 410，其中：

向量计算模块 402，用于计算目标文本中每个文本字符的字向量及每个文本分词的词向量。

向量拼接模块 404，用于将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字符的拼接向量。

第一文本特征获取模块 406，用于按照文本字符在目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到目标文本基于正向出现顺序的第一文本特征。

第二文本特征获取模块 408，用于按照文本字符在目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到目标文本基于逆向出现顺序的第二文本特征。

语义类型获取模块 410，用于将由第一文本特征与第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到目标文本的语义类型。

在一个实施例中，第一神经网络是按照文本字符在目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层，将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层，将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到目标文本基于正向出现顺序的第一文本特征。

在一个实施例中，如图 5 所示，上述还包括预设样本生成模块 412，基于预训练的第一神经网络层提取样本文本的字向量及词向量；对字向量和词向量分别进行字符编号；将字向量、词向量以及分别对应的字符编号写入到预设文件；计算文本字符的字向量和文本分词对应的词向量包括：对每个文本字符以及文本分词进行字符编号；基于字符编号，在预设文件中读取得到每个文本字符对应的字向量以及每个文本分词对应的词向量。

在一个实施例中，上述第一文本特征获取模块还用于按照文本字符在目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层；将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层；将下一隐层作为当前隐层，返回将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层的步骤，直至最后顺序文本字符，得到目标文本基于正向出现顺序的第一文本特征。

在一个实施例中，上述第二文本特征获取模块还用于按照文本字符在目标文本的逆向出现顺序，将第一文本特征、第一顺序文本字符对应的字向量和拼接向量输入至第二神经网络的第一隐层；将第一隐层的输出、当前顺序文本字符对应的字向量及拼接向量输入第二神经网络的当前隐层；将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第二神经网络的下一隐层；将下一隐层作为当前隐层进行，返回将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第二神经网络的下一隐层的步骤，直至最后顺序文本字符，得到目标文本基于逆向出现顺序的第二文本特征。

在一个实施例中，上述第一文本特征获取模块还用于将字向量和上一隐层的输出作为第一子隐层的输入，第一子隐层用于根据第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；将第一子字符特征和拼接向量作为第二子隐层的输入，第二子隐层用于根据第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

在一个实施例中，上述第一文本特征获取模块还用于将第一文本特征作为随机失活层的输入，随机失活层用于将第一文本特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征向量，作为第一神经网络的输出。

在一个实施例中，上述语义类型获取模块还用于将第一文本特征和第二文本特征进行拼接，得到目标文本的综合文本特征；将综合文本特征作为注意力机制层的输入，注意力机制层用于将综合文本特征中每个数据进行加权处理得到加权特征；将加权特征作为随机失活层的输入，随机失活层用于将加权特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征；将稀疏特征作为全连接层的输入，全连接层用于对稀疏特征进行分类运算得到每个语义类型对应的预测概率；选取预测概率最大的语义类型作为目标文本的语义类型。

上述实施例中，通过计算获得每个文本字符对应的字向量和所属文本分词的词向量并进行向量拼接，得到与文本字符对应的拼接向量，通过对文本字符进行向量拼接，通过多种特征向量来表征文本，增强了文本语言表示的特征维度。进一步，通过将字向量和拼接向量按照正向顺序和逆向顺序输入至不同神经网络的不同隐层，可以更充分的获取文本字符的相关信息，挖掘文本字符之间的上下文语义，使得经第一神经网络输出的第一文本特征和第二文本特征进行拼接得到综合特征，可以更充分的表达目标文本的语义特征，提高了文本语义识别的准确率。

关于文本语义识别装置的具体限定可以参见上文中对于文本语义识别方法的限定，在此不再赘述。上述文本语义识别装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图6所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储预设文件以及目标文本所包含的文本字符对应的字向量和文本分词对应的词向量。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种文本语义识别方法。

本领域技术人员可以理解，图6中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

一种计算机设备，包括存储器和处理器，所述存储器存储有计算机程序，其特征在于，所述处理器执行所述计算机程序时实现本申请任意一个实施例中提供的文本语义识别方法的步骤。

一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现本申请任意一个实施例中提供的文本语义识别方法的步骤。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，的计算机程序可存储于计算机可读存储介质中，其中，所述计算机可读存储介质可以是非易失性，也可以是易失性的。该计算机程序在执

行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器 (ROM)、可编程 ROM (PROM)、电可编程 ROM (EPROM)、电可擦除可编程 ROM (EEPROM) 或闪存。易失性存储器可包括随机存取存储器 (RAM) 或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM (SRAM)、动态 RAM (DRAM)、同步 DRAM (SDRAM)、双数据率 SDRAM (DDRSDRAM)、增强型 SDRAM (ESDRAM)、同步链路 (SynchLink) DRAM (SLDRAM)、存储器总线 (Rambus) 直接 RAM (RDRAM)、直接存储器总线动态 RAM (DRDRAM)、以及存储器总线动态 RAM (RDRAM) 等。

以上所述，仅为本申请的具体实施方式，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。



## 权利要求

1、一种文本语义识别方法，所述方法包括：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

2、根据权利要求1所述的方法，其中，所述方法还包括：

获取样本文本；

基于预训练的第一神经网络层提取所述样本文本的字向量及词向量；

对所述字向量和词向量分别进行字符编号；

将所述字向量、词向量以及分别对应的字符编号写入到预设文件；

所述计算文本字符的字向量和文本分词对应的词向量包括：

对每个所述文本字符以及所述文本分词进行字符编号；

基于所述字符编号，在所述预设文件中读取得到每个文本字符对应的字向量以及每个文本分词对应的词向量。

3、根据权利要求1所述的方法，其中，所述按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征包括：

按照文本字符在所述目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层；

将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层；

将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于正向出现顺序的第一文本特征。

4、根据权利要求3所述的方法，其中，所述当前隐层包括第一子隐层和第二子隐层；所述将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层包括：

将所述字向量和上一隐层的输出作为第一子隐层的输入，所述第一子隐层用于根据所述第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；

将所述第一子字符特征和所述拼接向量作为第二子隐层的输入，所述第二子隐层用于根据所述第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

5、根据权利要求3所述的方法，其中，所述第一神经网络还包括随机失活层，所述方法还包括：

将所述第一文本特征作为所述随机失活层的输入，所述随机失活层用于将所述第一文本特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征向量，作为所述第一神经网络的输出。

6、根据权利要求1所述的方法，其中，所述按照文本字符在所述目标文本的逆向出现

顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征包括：

按照文本字符在所述目标文本的逆向出现顺序，将所述第一文本特征、第一顺序文本字符对应的字向量和拼接向量输入至第二神经网络的第一隐层，得到所述第一隐层输出的字符特征；

将第二神经网络的第二隐层作为当前隐层，将第二顺序文本字符作为当前顺序文本字符；

将所述前一隐层输出的字符特征、当前顺序文本字符对应的字向量及拼接向量输入第二神经网络的当前隐层；

将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于逆向出现顺序的第二文本特征。

7、根据权利要求1所述的方法，其中，所述第三神经网络层包括注意力机制层和全连接层；将所述第一文本特征和所述第二文本特征进行拼接输入至第三神经网络，得到所述目标文本的语义类型包括：

将所述第一文本特征和所述第二文本特征进行拼接，得到所述目标文本的综合文本特征；

将所述综合文本特征作为所述注意力机制层的输入，所述注意力机制层用于将综合文本特征中每个数据进行加权处理得到加权特征；

将所述加权特征作为所述随机失活层的输入，所述随机失活层用于将加权特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征；

将所述稀疏特征作为全连接层的输入，所述全连接层用于对所述稀疏特征进行分类运算得到每个语义类型对应的预测概率；

选取所述预测概率最大的语义类型作为目标文本的语义类型。

8、一种文本语义识别装置，其中，所述装置包括：

向量计算模块，用于计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

向量拼接模块，用于将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

第一文本特征获取模块，用于按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

第二文本特征获取模块，用于按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

语义类型获取模块，用于将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

9、一种计算机设备，包括存储器和处理器，所述处理器、和所述存储器相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器用于执行所述存储器的所述程序指令，其中：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

10、根据权利要求9所述的计算机设备，其中，所述处理器用于：

获取样本文本；

基于预训练的第一神经网络层提取所述样本文本的字向量及词向量；

对所述字向量和词向量分别进行字符编号；

将所述字向量、词向量以及分别对应的字符编号写入到预设文件；

所述计算文本字符的字向量和文本分词对应的词向量包括：

对每个所述文本字符以及所述文本分词进行字符编号；

基于所述字符编号，在所述预设文件中读取得到每个文本字符对应的字向量以及每个文本分词对应的词向量。

11、根据权利要求9所述的计算机设备，其中，所述处理器用于：

按照文本字符在所述目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层；

将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层；

将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于正向出现顺序的第一文本特征。

12、根据权利要求11所述的计算机设备，其中，所述当前隐层包括第一子隐层和第二子隐层；所述处理器用于：

将所述字向量和上一隐层的输出作为第一子隐层的输入，所述第一子隐层用于根据所述第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；

将所述第一子字符特征和所述拼接向量作为第二子隐层的输入，所述第二子隐层用于根据所述第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

13、根据权利要求11所述的计算机设备，其中，所述处理器用于：

将所述第一文本特征作为所述随机失活层的输入，所述随机失活层用于将第一文本特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征向量，作为所述第一神经网络的输出。

14、根据权利要求9所述的计算机设备，其中，所述处理器用于：

按照文本字符在所述目标文本的逆向出现顺序，将所述第一文本特征、第一顺序文本字符对应的字向量和拼接向量输入至第二神经网络的第一隐层，得到所述第一隐层输出的字符特征；

将第二神经网络的第二隐层作为当前隐层，将第二顺序文本字符作为当前顺序文本字符；

将所述前一隐层输出的字符特征、当前顺序文本字符对应的字向量及拼接向量输入第二神经网络的当前隐层；

将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于逆向出现顺序的第二文本特征。

15、根据权利要求9所述的计算机设备，其中，所述第三神经网络层包括注意力机制层和全连接层；所述处理器用于：

将所述第一文本特征和所述第二文本特征进行拼接，得到所述目标文本的综合文本特征；

将所述综合文本特征作为所述注意力机制层的输入，所述注意力机制层用于将综合文本特征中每个数据进行加权处理得到加权特征；

将所述加权特征作为所述随机失活层的输入，所述随机失活层用于将加权特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征；

将所述稀疏特征作为全连接层的输入，所述全连接层用于对所述稀疏特征进行分类运算得到每个语义类型对应的预测概率；

选取所述预测概率最大的语义类型作为目标文本的语义类型。

16、一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令被处理器执行时，用于实现以下步骤：

计算目标文本中每个文本字符的字向量及每个文本分词的词向量；

将每个文本字符的字向量与所属文本分词的词向量进行拼接，得到相应文本字符的拼接向量；

按照文本字符在所述目标文本的正向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第一神经网络的不同隐层，得到所述目标文本基于正向出现顺序的第一文本特征；

按照文本字符在所述目标文本的逆向出现顺序，将多个文本字符对应的字向量及拼接向量依次输入第二神经网络的不同隐层，得到所述目标文本基于逆向出现顺序的第二文本特征；

将由所述第一文本特征与所述第二文本特征拼接得到的综合文本特征输入至第三神经网络，得到所述目标文本的语义类型。

17、根据权利要求16所述的计算机可读存储介质，其中，所述程序指令被处理器执行时，还用于实现以下步骤：

获取样本文本；

基于预训练的第一神经网络层提取所述样本文本的字向量及词向量；

对所述字向量和词向量分别进行字符编号；

将所述字向量、词向量以及分别对应的字符编号写入到预设文件；

所述计算文本字符的字向量和文本分词对应的词向量包括：

对每个所述文本字符以及所述文本分词进行字符编号；

基于所述字符编号，在所述预设文件中读取得到每个文本字符对应的字向量以及每个文本分词对应的词向量。

18、根据权利要求16所述的计算机可读存储介质，其中，所述程序指令被处理器执行时，还用于实现以下步骤：

按照文本字符在所述目标文本的正向出现顺序，将当前顺序文本字符对应的字向量及拼接向量输入第一神经网络的当前隐层；

将当前隐层输出的字符特征与下一顺序文本字符对应的字向量及拼接向量输入第一神经网络的下一隐层；

将下一隐层作为当前隐层进行迭代，直至最后顺序文本字符，得到所述目标文本基于正向出现顺序的第一文本特征。

19、根据权利要求18所述的计算机可读存储介质，其中，所述当前隐层包括第一子隐层和第二子隐层；所述程序指令被处理器执行时，还用于实现以下步骤方法：

将所述字向量和上一隐层的输出作为第一子隐层的输入，所述第一子隐层用于根据所述第一子隐层对应的各个神经元节点的权重对字向量进行投影得到第一子字符特征；

将所述第一子字符特征和所述拼接向量作为第二子隐层的输入，所述第二子隐层用于根据所述第二子隐层对应的各个神经元节点的权重对拼接向量进行投影得到第二子字符特征，作为当前隐层的输出。

20、根据权利要求 18 所述的计算机可读存储介质，其中，所述程序指令被处理器执行时，还用于实现以下步骤：

将所述第一文本特征作为所述随机失活层的输入，所述随机失活层用于将第一文本特征中每个数据按照预设的稀疏概率进行投影得到稀疏特征向量，作为所述第一神经网络的输出。

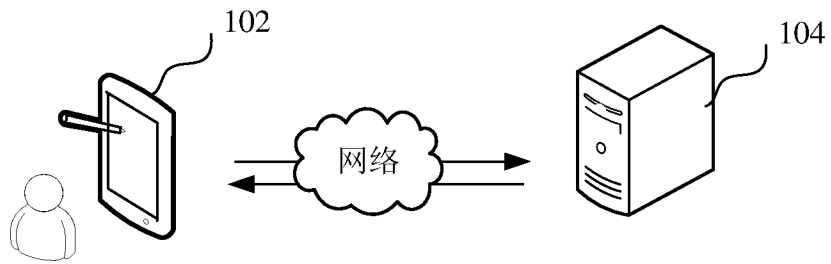


图 1

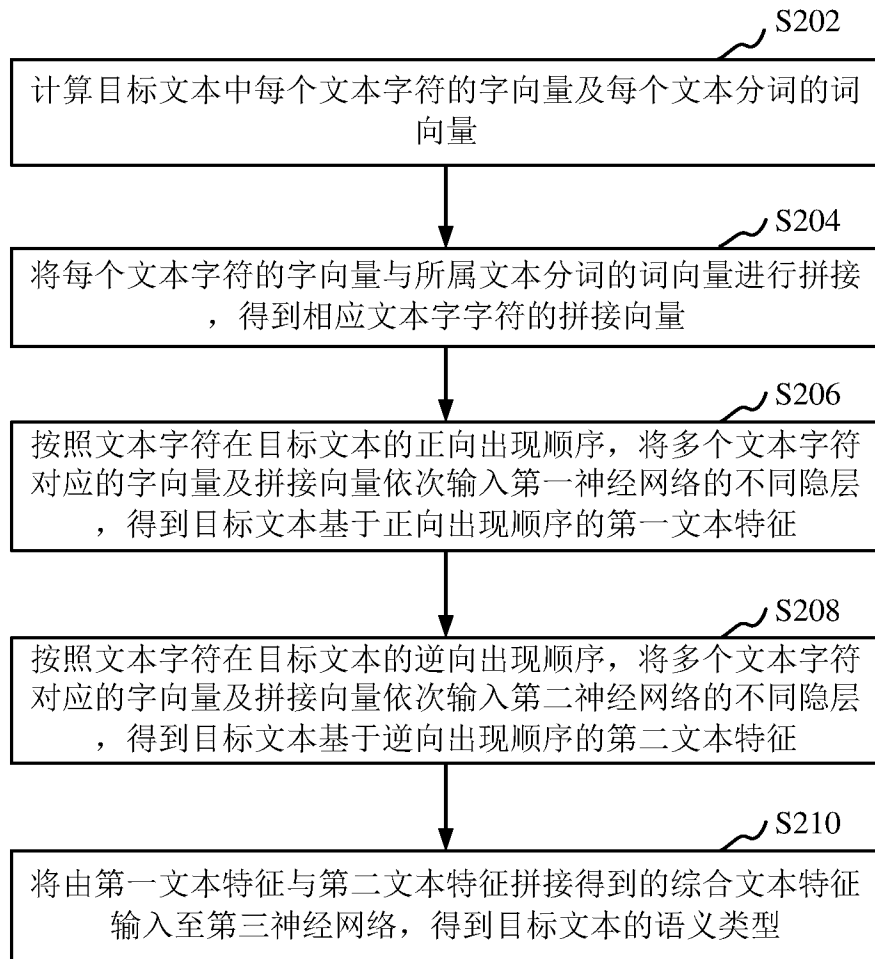


图 2

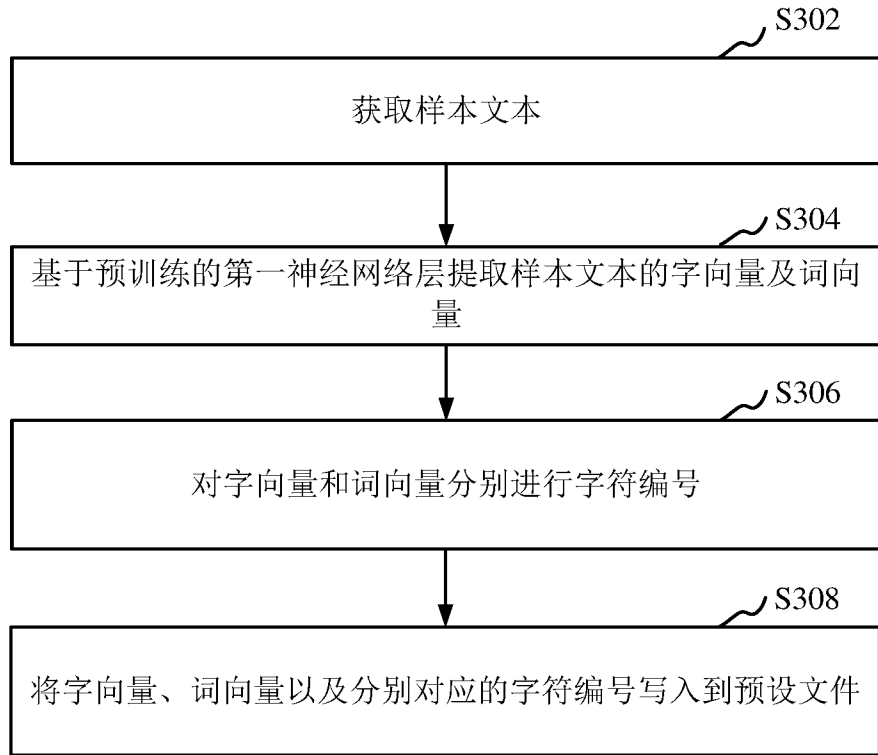


图 3

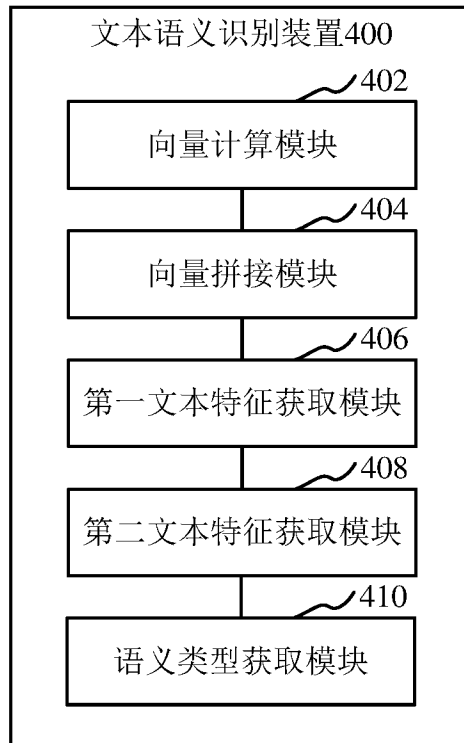


图 4

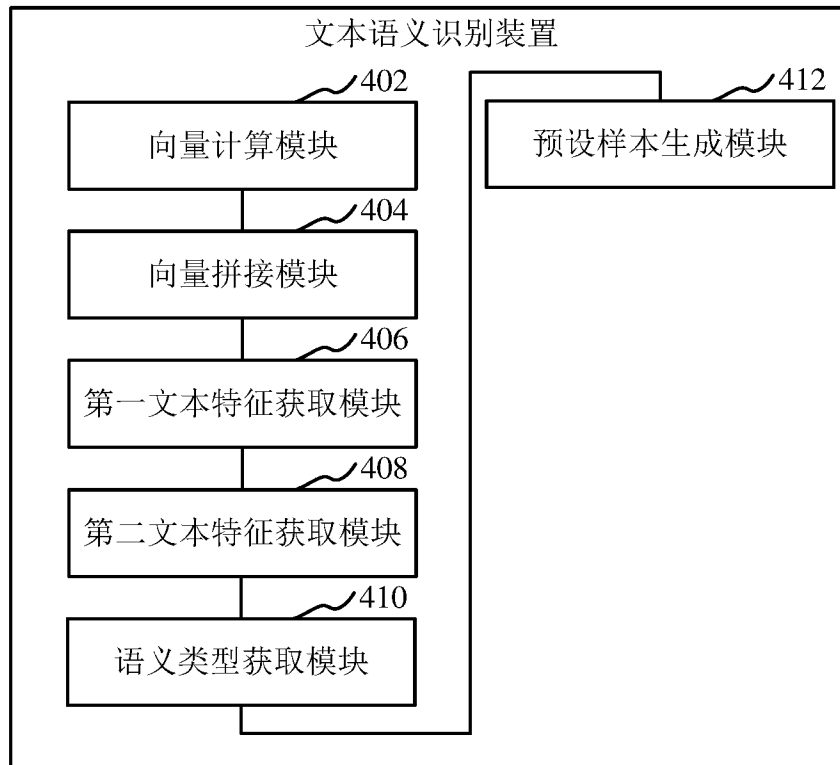


图 5

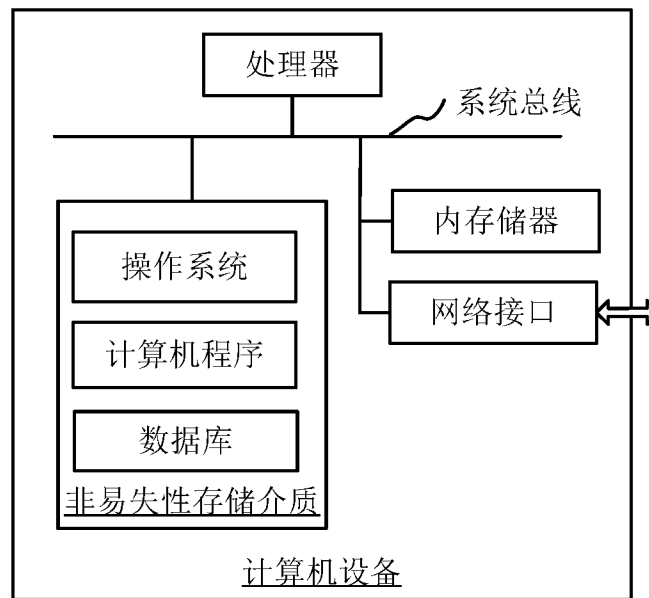


图 6



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/104679

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> G06F 17/00(2019.01)i  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06F  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) DWPI; USTXT; VEN; CNABS; CNTXT; EPTXT; WOTXT; CNKI; GOOGLE: 文本, 文档, 语义, 分类, 类型, 类别, 词向量, 字向量, 拼接, 正向, 逆向, 神经网络, text, document, semantic, classify, type, class, word vector, character vector, concatenate, forward, backward, neural network		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110598206 A (PINGAN INTERNATIONAL WISDOM CITY TECHNOLOGY CO., LTD.) 20 December 2019 (2019-12-20) description, paragraphs 58-155	1-20
A	CN 109684626 A (DEEP THOUGHTS ARTIFICIAL INTELLIGENCE ROBOT TECHNOLOGY (BEIJING) CO., LTD.) 26 April 2019 (2019-04-26) description, paragraphs 28-71	1-20
A	CN 108132931 A (BEIJING ULTRAPOWERSOFTWARE CO., LTD.) 08 June 2018 (2018-06-08) description, paragraphs 98-170	1-20
A	CN 109376240 A (PING AN TECHNOLOGY SHENZHEN CO., LTD.) 22 February 2019 (2019-02-22) entire document	1-20
A	US 2018018573 A1 (XEROX CORPORATION) 18 January 2018 (2018-01-18) entire document	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search <b>10 October 2020</b>		Date of mailing of the international search report <b>22 October 2020</b>
Name and mailing address of the ISA/CN <b>China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088 China</b> Facsimile No. (86-10)62019451		Authorized officer  Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2020/104679**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110598206	A	20 December 2019	None			
CN	109684626	A	26 April 2019	None			
CN	108132931	A	08 June 2018	None			
CN	109376240	A	22 February 2019	WO	2020073673	A1	16 April 2020
US	2018018573	A1	18 January 2018	None			

<p><b>A. 主题的分类</b></p> <p>G06F 17/00 (2019.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																				
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>DWPI;USTXT;VEN;CNABS;CNTXT;EPTXT;WOTXT;CNKI;GOOGLE:文本, 文档, 语义, 分类, 类型, 类别, 词向量, 字向量, 拼接, 正向, 逆向, 神经网络, text, document, semantic, classify, type, class, word vector, character vector, concatenate, forward, backward, neural network</p>																				
<p><b>C. 相关文件</b></p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 110598206 A (平安国际智慧城市科技股份有限公司) 2019年 12月 20日 (2019 - 12 - 20) 说明书58-155段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 109684626 A (深思考人工智能机器人科技北京有限公司) 2019年 4月 26日 (2019 - 04 - 26) 说明书28-71段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 108132931 A (北京神州泰岳软件股份有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书98-170段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 109376240 A (平安科技深圳有限公司) 2019年 2月 22日 (2019 - 02 - 22) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2018018573 A1 (XEROX CORPORATION) 2018年 1月 18日 (2018 - 01 - 18) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 110598206 A (平安国际智慧城市科技股份有限公司) 2019年 12月 20日 (2019 - 12 - 20) 说明书58-155段	1-20	A	CN 109684626 A (深思考人工智能机器人科技北京有限公司) 2019年 4月 26日 (2019 - 04 - 26) 说明书28-71段	1-20	A	CN 108132931 A (北京神州泰岳软件股份有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书98-170段	1-20	A	CN 109376240 A (平安科技深圳有限公司) 2019年 2月 22日 (2019 - 02 - 22) 全文	1-20	A	US 2018018573 A1 (XEROX CORPORATION) 2018年 1月 18日 (2018 - 01 - 18) 全文	1-20
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 110598206 A (平安国际智慧城市科技股份有限公司) 2019年 12月 20日 (2019 - 12 - 20) 说明书58-155段	1-20																		
A	CN 109684626 A (深思考人工智能机器人科技北京有限公司) 2019年 4月 26日 (2019 - 04 - 26) 说明书28-71段	1-20																		
A	CN 108132931 A (北京神州泰岳软件股份有限公司) 2018年 6月 8日 (2018 - 06 - 08) 说明书98-170段	1-20																		
A	CN 109376240 A (平安科技深圳有限公司) 2019年 2月 22日 (2019 - 02 - 22) 全文	1-20																		
A	US 2018018573 A1 (XEROX CORPORATION) 2018年 1月 18日 (2018 - 01 - 18) 全文	1-20																		
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&amp;” 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2020年 10月 10日</p>	<p>国际检索报告邮寄日期</p> <p>2020年 10月 22日</p>																			
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>	<p>授权官员</p> <p>薛双双</p> <p>电话号码 86-(20)-28958967</p>																			

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2020/104679

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	110598206	A	2019年 12月 20日	无			
CN	109684626	A	2019年 4月 26日	无			
CN	108132931	A	2018年 6月 8日	无			
CN	109376240	A	2019年 2月 22日	WO	2020073673	A1	2020年 4月 16日
US	2018018573	A1	2018年 1月 18日	无			