



(12) 发明专利申请

(10) 申请公布号 CN 115936003 A

(43) 申请公布日 2023.04.07

(21) 申请号 202211522902.2

G06N 3/08 (2023.01)

(22) 申请日 2022.11.30

(71) 申请人 湖南科创信息技术股份有限公司
地址 410009 湖南省长沙市岳麓区青山路
678号科创软件园

(72) 发明人 虞凌云 李海超 刘芝冰

(74) 专利代理机构 长沙智嵘专利代理事务所
(普通合伙) 43211

专利代理师 黄海波

(51) Int. Cl.

G06F 40/30 (2020.01)

G06F 40/194 (2020.01)

G06F 18/214 (2023.01)

G06F 18/24 (2023.01)

G06N 3/0464 (2023.01)

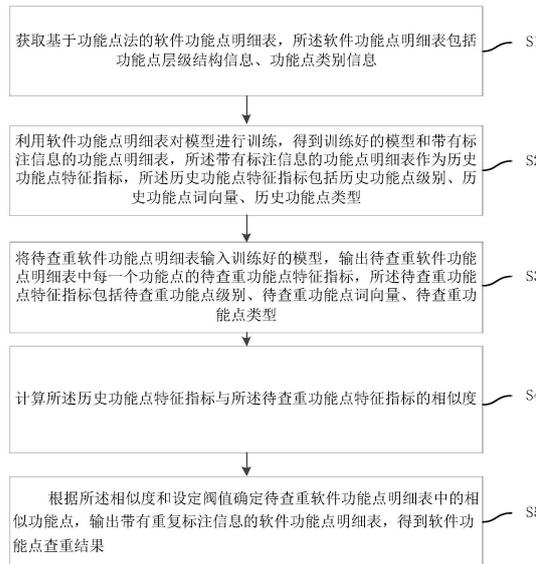
权利要求书3页 说明书10页 附图8页

(54) 发明名称

基于神经网络的软件功能点查重方法、装置、设备及介质

(57) 摘要

本申请公开了一种基于神经网络的软件功能点查重方法、装置、设备及介质,所述方法包括步骤:获取基于功能点法的软件功能点明细表;利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标;将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标;计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度;根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表。本申请提高了查重的效率和准确率,减少了项目重复建设成本。



1. 一种基于神经网络的软件功能点查重方法,其特征在于,包括步骤:

获取基于功能点法的软件功能点明细表,所述软件功能点明细表包括功能点层级结构信息、功能点类别信息;

利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的软件功能点明细表,所述带有标注信息的软件功能点明细表作为历史功能点特征指标,所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型;

将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型;

计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度;

根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表,得到软件功能点查重结果。

2. 根据权利要求1所述的基于神经网络的软件功能点查重方法,其特征在于,所述利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的软件功能点明细表,所述带有标注信息的软件功能点明细表作为历史功能点特征指标,具体包括步骤:

通过功能点分级模型将软件功能点明细表中的层级结构读取出来形成功能点文本,将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据;

采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义;

以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算。

3. 根据权利要求2所述的基于神经网络的软件功能点查重方法,其特征在于,所述采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义,具体包括步骤:

训练语料构建:从包括补全后的功能点文本数据与海量的中文数据的完全没有标注信息的文本中构建用于词向量模型的自训练语料;

词向量模型构建:基于CBOW构建词向量模型,所述词向量模型利用文本上下文推理出缺失的真实词,获取到更加准确的文本语义,并将词转化成代表文本语义的向量,语义越相似的词所代表的词向量距离越接近;

词向量模型训练:使用准备好的自训练语料对词向量模型进行训练,使词向量模型具备语义理解能力;

词向量模型保存:将训练好的词向量模型进行保存,用于直接对功能点词向量进行计算。

4. 根据权利要求2所述的基于神经网络的软件功能点查重方法,其特征在于,所述以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算,具体包括步骤:

构建训练语料:以补全后的功能点文本数据作为训练语料,取功能点类别作为标签,保证训练数据的准确性;

构建功能点分类模型:功能点分类模型采用卷积神经网络模型,包括输入层、卷积层、池化层、全连接层和输出层;

训练功能点分类模型:使用准备好的带有类别信息的功能点文本数据经过预处理过程对功能点分类模型进行训练,通过反向传播技术调整模型参数,并经过多次调整超参数保证模型训练到最优的状态,最终正确的对功能点文本数据进行分类;

输出分类结果:将训练好的功能点分类模型进行保存,用于对软件功能点进行分类。

5. 根据权利要求1所述的基于神经网络的软件功能点查重方法,其特征在于,所述将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,具体包括步骤:

功能点分级模型将待查重软件功能点明细表中的层级结构读取出来形成功能点文本,再将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据后,输出并保存待查重功能点级别;

功能点词向量模型利用所述功能点文本数据上下文推理出缺失的真实词,获取到更加准确的文本语义,将词转化成代表文本语义的向量后,输出并保存待查重功能点词向量;

功能点分类模型通过提取功能点文本数据的关键语义特征,识别文本之间的隐藏关系,正确的对功能点文本进行分类后,输出并保存待查重软件功能点明细表中每一个待查重功能点类型。

6. 根据权利要求1所述的基于神经网络的软件功能点查重方法,其特征在于,所述计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度,具体包括步骤:

分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度;

根据待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度加权计算待查重功能点相似度:

$$W = \alpha * X + \beta * Y + \theta * Z$$

式中:

α 、 β 、 θ 为对应指标所占权重;

W为待查重功能点相似度;

X为待查重功能点级别相似度;

Y为待查重功能点词向量相似度;

Z为待查重功能点类型相似度。

7. 根据权利要求6所述的基于神经网络的软件功能点查重方法,其特征在于,所述分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度,具体包括步骤:

计算待查重功能点级别相似度时,采用逐级对比原则,根据待查重功能点在功能模块中的级别深度确定待查重功能点级别相似度,其中,所述待查重功能点级别相似度的大小与待查重功能点在功能模块中的级别深度正相关;

计算待查重功能点词向量相似度时,通过将功能点文本数据转换成词向量,计算历史

功能点词向量和待查重功能点词向量的余弦相似度得到二者之间的相似度；

计算待查重功能点类型相似度时，将模型输出的功能点类型分为EI、EQ、EO、ELF、ILF五类，根据功能点类型的不同，设置不同功能点之间的相似度。

8. 一种基于神经网络的软件功能点查重装置，其特征在于，包括：

功能点明细表获取模块，用于获取基于功能点法的软件功能点明细表，所述软件功能点明细表包括功能点层级结构信息、功能点类别信息；

历史功能点特征指标计算模块，用于利用软件功能点明细表对模型进行训练，得到训练好的模型和带有标注信息的功能点明细表，所述带有标注信息的功能点明细表作为历史功能点特征指标，所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型；

待查重功能点特征指标计算模块，用于将待查重软件功能点明细表输入训练好的模型，输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标，所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型；

相似度计算模块，用于计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度；

功能点查重结果输出模块，用于根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点，输出带有重复标注信息的软件功能点明细表，得到软件功能点查重结果。

9. 一种电子设备，包括存储器、处理器及存储在存储器上并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如权利要求1至7中任一项所述基于神经网络的软件功能点查重方法的步骤。

10. 一种存储介质，所述存储介质包括存储的程序，其特征在于，

在所述程序运行时控制所述存储介质所在的设备执行如权利要求1至7中任一项所述基于神经网络的软件功能点查重方法的步骤。

基于神经网络的软件功能点查重方法、装置、设备及介质

技术领域

[0001] 本申请涉及计算机技术领域,特别地,涉及一种基于神经网络的软件功能点查重方法、装置、设备及介质。

背景技术

[0002] 功能点法(IFPUG-FPA)是一种基于用户需求、有具体算法支撑、客观独立于技术实现、且可在实际开发前进行估算的一种软件规模度量方法,具有强科学性、权威性和可操作性等特点,是国家标准、行业标准采用的方法。功能点计数过程是指量化功能用户需求和评价非功能点用户需求,以对应应用系统软件功能计数的过程。

[0003] 考虑软件项目越来越多,在同一领域存在重复或类似项目、同一项目中存在相同或类似功能点等,通过对功能点重复性检查,可发现不合理功能项,有效避免重复建设。但积累了大量项目的软件功能点,仅靠人工查重难以进行,人工成本和等待成本必然巨大,同时需要人工逐条识别、判断功能点数据,难以克服主观性和生理疲劳等因素,很难保证查重质量。

[0004] 因此如何有效减少用户对功能点查重时间和成本,提高查重准确率是技术人员亟需解决的问题。

发明内容

[0005] 本申请一方面提供了一种基于神经网络的软件功能点查重方法,以解决现有技术功能点重复性检查时时间和成本高、效率和准确率低的技术问题。

[0006] 本申请采用的技术方案如下:

[0007] 一种基于神经网络的软件功能点查重方法,包括步骤:

[0008] 获取基于功能点法的软件功能点明细表,所述软件功能点明细表包括功能点层级结构信息、功能点类别信息;

[0009] 利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型;

[0010] 将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型;

[0011] 计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度;

[0012] 根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表,得到软件功能点查重结果。

[0013] 优选地,所述利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,具体包括步骤:

[0014] 通过功能点分级模型将软件功能点明细表中的层级结构读取出来形成功能点文本,将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据;

[0015] 采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义;

[0016] 以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算。

[0017] 优选地,所述采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义,具体包括步骤:

[0018] 训练语料构建:从包括补全后的功能点文本数据与海量的中文数据的完全没有标注信息的文本中构建用于词向量模型的自训练语料;

[0019] 词向量模型构建:基于CBOW构建词向量模型,所述词向量模型利用文本上下文推理出缺失的真实词,获取到更加准确的文本语义,并将词转化成代表文本语义的向量,语义越相似的词所代表的词向量距离越接近;

[0020] 词向量模型训练:使用准备好的自训练语料对词向量模型进行训练,使词向量模型具备语义理解能力;

[0021] 词向量模型保存:将训练好的词向量模型进行保存,用于直接对功能点词向量进行计算。

[0022] 优选地,所述以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算,具体包括步骤:

[0023] 构建训练语料:以补全后的功能点文本数据作为训练语料,取功能点类别作为标签,保证训练数据的准确性;

[0024] 构建功能点分类模型:功能点分类模型采用卷积神经网络模型,包括输入层、卷积层、池化层、全连接层和输出层;

[0025] 训练功能点分类模型:使用准备好的带有类别信息的功能点文本数据经过预处理过程对功能点分类模型进行训练,通过反向传播技术调整模型参数,并经过多次调整超参数保证模型训练到最优的状态,最终正确的对功能点文本数据进行分类;

[0026] 输出分类结果:将训练好的功能点分类模型进行保存,用于对软件功能点进行分

类。

[0027] 优选地,所述将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,具体包括步骤:

[0028] 功能点分级模型将待查重软件功能点明细表中的层级结构读取出来形成功能点文本,再将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据后,输出并保存待查重功能点级别;

[0029] 功能点词向量模型利用所述功能点文本数据上下文推理出缺失的真实词,获取到更加准确的文本语义,将词转化成代表文本语义的向量后,输出并保存待查重功能点词向量;

[0030] 功能点分类模型通过提取功能点文本数据的关键语义特征,识别文本之间的隐藏关系,正确的对功能点文本进行分类后,输出并保存待查重软件功能点明细表中每一个待查重功能点类型。

[0031] 优选地,所述计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度,具体包括步骤:

[0032] 分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度;

[0033] 根据待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度加权计算待查重功能点相似度:

[0034] $W = \alpha * X + \beta * Y + \theta * Z$

[0035] 式中:

[0036] α 、 β 、 θ 为对应指标所占权重;

[0037] W为待查重功能点相似度;

[0038] X为待查重功能点级别相似度;

[0039] Y为待查重功能点词向量相似度;

[0040] Z为待查重功能点类型相似度。

[0041] 优选地,所述分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度,具体包括步骤:

[0042] 计算待查重功能点级别相似度时,采用逐级对比原则,根据待查重功能点在功能模块中的级别深度确定待查重功能点级别相似度,其中,所述待查重功能点级别相似度的大小与待查重功能点在功能模块中的级别深度正相关;

[0043] 计算待查重功能点词向量相似度时,通过将功能点文本数据转换成词向量,计算历史功能点词向量和待查重功能点词向量的余弦相似度得到二者之间的相似度;

[0044] 计算待查重功能点类型相似度时,将模型输出的功能点类型分为EI、EQ、EO、ELF、ILF五类,根据功能点类型的不同,设置不同功能点之间的相似度。

[0045] 本申请另一方面还提供了一种基于神经网络的软件功能点查重装置,包括:

[0046] 功能点明细表获取模块,用于获取基于功能点法的软件功能点明细表,所述软件功能点明细表包括功能点层级结构信息、功能点类别信息;

[0047] 历史功能点特征指标计算模块,用于利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型;

[0048] 待查重功能点特征指标计算模块,用于将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型;

[0049] 相似度计算模块,用于计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度;

[0050] 功能点查重结果输出模块,用于根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表,得到软件功能点

查重结果。

[0051] 本申请另一方面还提供了一种电子设备,包括存储器、处理器及存储在存储器上并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所述基于神经网络的软件功能点查重方法的步骤。

[0052] 本申请另一方面还提供了一种存储介质,所述存储介质包括存储的程序,在所述程序运行时控制所述存储介质所在的设备执行所述基于神经网络的软件功能点查重方法的步骤。

[0053] 相比现有技术,本申请具有以下有益效果:

[0054] 本申请提供的通过机器方式代替人工方式进行功能点查重,提高了查重的效率和准确率,减少了项目重复建设,具体地,本申请采用CBOW空间语义向量模型、卷积神经网络(CNN)分类模型,基于深度语义的词向量生成算法,对功能点进行正确的分类,将功能点文本转换为具有上下文语义特征的词向量和正确的类别,使模型具备了与业务专家相近水准的文本语义理解能力,同时结合功能点类型和分级,能极大提高查重的准确率。同时,本申请受大数据聚类技术启发,基于功能点文本转换为具有上下文语义特征的词向量,支持海量功能点文本比较的快速索引算法,通过该方法能够快速比较功能点文本之间的相似性,将查重需要的计算量减少为原来的1/100,保证了在0.1秒内完成10万条历史需求的查重,极大的提高了查重的效率,能在短时间内快速完成海量数据的查重。本申请计算功能点相似度采用多个维度指标对比,再采用加权方式计算功能点相似度,相比采用单一维度词向量余弦相似度计算功能点文本相似度,增加了功能点层级和类型比较,有效的提高查重的准确率。

[0055] 除了上面所描述的目的、特征和优点之外,本申请还有其它的目的、特征和优点。下面将参照附图,对本申请作进一步详细的说明。

附图说明

[0056] 构成本申请的一部分的附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0057] 图1是本申请优选实施例的基于神经网络的软件功能点查重方法流程示意图。

[0058] 图2是本申请优选实施例的软件功能点明细表示意图。

[0059] 图3是本申请优选实施例的步骤S2的子步骤流程示意图。

[0060] 图4是本申请优选实施例的步骤S22的子步骤流程示意图。

[0061] 图5是本申请优选实施例的步骤S23的子步骤流程示意图。

[0062] 图6是本申请优选实施例的步骤S3的子步骤流程示意图。

[0063] 图7是本申请优选实施例的步骤S4的子步骤流程示意图。

[0064] 图8是本申请优选实施例的步骤S41的子步骤流程示意图。

[0065] 图9是本申请优选实施例的待查重功能点级别相似度计算流程示意图。

[0066] 图10是本申请优选实施例的基于神经网络的软件功能点查重装置模块示意图。

[0067] 图11是本申请优选实施例的电子设备实体示意框图。

[0068] 图12是本申请优选实施例的计算机设备的内部结构图。

具体实施方式

[0069] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0070] 参照图1,本申请的优选实施例提供了一种基于神经网络的软件功能点查重方法,包括步骤:

[0071] S1、获取基于功能点法的软件功能点明细表,所述软件功能点明细表包括功能点层级结构信息、功能点类别信息,如图2所示,软件功能点明细表为企业在使用功能点法评估项目工作量过程中产生,经过沉淀和积累形成。软件功能点明细表清晰记录功能点的层级结构、功能点类型等信息,该数据因经由多次评估和审核,逐层质量把关,数据质量和可靠性可充分保证;

[0072] S2、利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型;

[0073] S3、将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型;

[0074] S4、计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度;

[0075] S5、根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表,得到软件功能点查重结果。

[0076] 本实施例提供的通过机器方式代替人工方式进行功能点查重,提高了查重的效率和准确率,减少了项目重复建设,具体地,本申请采用CBOW空间语义向量模型、卷积神经网络(CNN)分类模型,基于深度语义的词向量生成算法,对功能点进行正确的分类,将功能点文本转换为具有上下文语义特征的词向量和正确的类别,使模型具备了与业务专家相近水准的文本语义理解能力,同时结合功能点类型和分级,能极大提高查重的准确率。同时,本实施例受大数据聚类技术启发,基于功能点文本转换为具有上下文语义特征的词向量,支持海量功能点文本比较的快速索引算法,通过该方法能够快速比较功能点文本之间的相似性,将查重需要的计算量减少为原来的1/100,保证了在0.1秒内完成10万条历史需求的查重,极大的提高了查重的效率,能在短时间内快速完成海量数据的查重。本实施例计算功能点相似度采用多个维度指标对比,再采用加权方式计算功能点相似度,相比采用单一维度词向量余弦相似度计算功能点文本相似度,增加了功能点层级和类型比较,有效的提高查重的准确率。

[0077] 优选地,如图3所示,所述利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,具体包括步骤:

[0078] S21、通过功能点分级模型将软件功能点明细表中的层级结构(包括一级、二级、三级、四级等)读取出来形成功能点文本,将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据,比如:原始的功能点描述为“登录功能点单点登录”,经过补全后的文本为“一级功能为登录功能,二级功能为单点登录”。通过功能点文本的信息补全,可以帮助后续模型理解功能点文本的层级信息,做出更加合理的判断;

[0079] 以图2中序号为1的功能点为例,机器读取到的功能点文本为“人脸基础信息管理人脸本地管理人脸基础管理人脸基础管理用户组创建”,经过补全后的文本为:“功能模块人脸基础信息管理,一级功能为人脸本地管理,二级功能为人脸基础管理,三级功能为人脸基础管理,四级功能为用户组创建”。

[0080] S22、采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义。词向量模型可以把词映射到高维空间,在高维空间中意思相近的词之间的距离彼此较近,这样使计算机模仿人读书识字的过程,从而具备语义理解的能力。为了方便后续训练,将上一步骤中完成分级的功能点文本按照8:2比例随机划分为训练数据、测试数据;

[0081] S23、以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算。

[0082] 优选地,如图4所示,所述采集补全后的功能点文本数据与海量的中文数据来训练词向量模型,使词向量模型能够理解功能点文本的真实语义,具体包括步骤:

[0083] S221、训练语料构建:从包括补全后的功能点文本数据与海量的中文数据的完全没有标注信息的文本中构建用于词向量模型的自训练语料;

[0084] S222、词向量模型构建:基于CBOW构建词向量模型,所述词向量模型利用文本上下文推理出缺失的真实词,获取到更加准确的文本语义,并将词转化成代表文本语义的向量,语义越相似的词所代表的词向量距离越接近;

[0085] S223、词向量模型训练:使用准备好的自训练语料对词向量模型进行训练,使词向量模型具备语义理解能力,因此需要使用大量的训练语料供给词向量模型进行训练,通过反向传播技术不断调整模型的参数,并经过多次调整超参数保证模型训练到最优的状态,最终使得词向量模型能够具备语义理解能力,可视两条功能点词向量在二维等高平面的距离关系;

[0086] S224、词向量模型保存:将训练好的词向量模型进行保存,用于直接对功能点词向量进行计算。

[0087] 优选地,如图5所示,所述以补全后的功能点文本数据作为训练语料、功能点类别作为训练标签,对模型进行训练,得到功能点分类模型用于直接提取功能点文本数据的关键语义特征,并基于所述关键语义特征进行功能点类别概率计算,具体包括步骤:

[0088] S231、构建训练语料:以补全后的功能点文本数据作为训练语料,取功能点类别作为标签,保证训练数据的准确性;

[0089] S232、构建功能点分类模型:功能点分类模型采用卷积神经网络模型,包括输入层、卷积层、池化层、全连接层和输出层,在卷积层以一定的间隔滑动滤波器的窗口进行卷积运算在所有位置都进行一遍,就可以得到卷积运算的输出;在池化层利用局部词序信息,提取初级的特征,并组合初级的特征为高级特征,通过卷积与池化操作,省去了传统机器学习中的特征工程的步骤;全连接层把卷积与池化层提取的特征输入到一个分类器中进行分类,输出层采用Softmax分类模型将功能点文本数据分为5个类别,分别对应功能点文本的“EI”、“EQ”、“EO”、“EIL”和“ILF”;

[0090] S233、训练功能点分类模型:使用准备好的带有类别信息的功能点文本数据经过

预处理过程后对功能点分类模型进行训练,通过反向传播技术调整模型参数,并经过多次调整超参数保证模型训练到最优的状态,最终正确的对功能点文本数据进行分类;

[0091] S234、输出分类结果:将训练好的功能点分类模型进行保存,用于对软件功能点进行

分类。
[0092] 优选地,如图6所示,所述将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,具体包括步骤:

[0093] S31、功能点分级模型将待查重软件功能点明细表中的层级结构读取出来形成功能点文本,再将功能点文本的层级结构通过自然语言描述出来,添加到功能点文本中,补全功能点文本数据后,输出并保存待查重功能点级别;

[0094] S32、功能点词向量模型利用所述功能点文本数据上下文推理出缺失的真实词,获取到更加准确的文本语义,将词转化成代表文本语义的向量后,输出并保存待查重功能点词向量;

[0095] S33、功能点分类模型通过提取功能点文本数据的关键语义特征,识别文本之间的隐藏关系,正确的对功能点文本进行分类后,输出并保存待查重软件功能点明细表中每一个待查重功能点类型。

[0096] 优选地,如图7所示,所述计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度,具体包括步骤:

[0097] S41、分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度;

[0098] 具体地,如图8所示,所述分别计算待查重功能点的待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度,具体包括步骤:

[0099] S411、如图9所示,计算待查重功能点级别相似度时,采用逐级对比原则,根据待查重功能点在功能模块中的级别深度确定待查重功能点级别相似度,其中,所述待查重功能点级别相似度的大小与待查重功能点在功能模块中的级别深度正相关,即功能点级别相似度(X)为表1所示:

[0100] 表1:

序号	相似内容	相似度 (X)
1	全部不同	0
2	功能模块相似	20%
3	一级功能相似	40%
4	二级功能相似	60%
5	三级功能相似	80%
6	四级功能相似	100%

[0103] 本实施例根据历史功能点文本级别和待查重功能点的级别查询表1,如待查重功

能点的类型与历史功能点文本级别在二级相似,则待查重功能点级别相似度X为60%。

[0104] S412、计算待查重功能点词向量相似度时,通过将功能点文本数据转换成词向量,计算历史功能点词向量A和待查重功能点词向量B的余弦相似度得到二者之间的相似度:

$$[0105] \quad Y = \frac{\sum_1^n A_i \times B_i}{\sqrt{\sum_1^n (A_i)^2} \times \sqrt{\sum_1^n (B_i)^2}} = \frac{A \times B}{|A| \times |B|}$$

[0106] 式中:n表示空间向量维度数,i表示第i维空间;

[0107] 本实施例根据历史功能点词向量和待查重功能点词向量使用上式进行计算,如历史功能点词向量为A (1、3、2),待查重功能点的词向量为B (2、4、5),则待查重功能点词向量相似度Y为:

$$[0108] \quad Y = \frac{A \times B}{|A| \times |B|} = 95.63\%$$

[0109] S413、计算待查重功能点类型相似度时,将模型输出的功能点类型分为EI、EQ、EO、ELF、ILF五类,根据功能点类型的不同,设置不同功能点之间的类型相似度(Z),如表2所示:

[0110] 表2:

相似度 (Z)	EI	EQ	EO	ELF	ILF
EI	100	70%	60%	10%	20%
EQ	70%	100%	90%	20%	30%
EO	60%	90%	100%	30%	40%
ELF	10%	20%	30%	100%	50%
ILF	20%	30%	40%	50%	100%

[0112] 本实施例根据历史功能点类型和待查重功能点类型查询表2,如历史功能点类型为EI、待查重功能点类型为EO,则查重功能点类型相似度Z为60%;

[0113] S42、根据待查重功能点级别相似度、待查重功能点词向量相似度、待查重功能点类型相似度加权计算待查重功能点相似度:

$$[0114] \quad W = \alpha * X + \beta * Y + \theta * Z = 30\% * 60\% + 40\% * 95.63\% + 30\% * 60\% = 74.25\%,$$

[0115] 式中:

[0116] α 、 β 、 θ 为对应指标所占权重,分别设置为:30%、40%、30%;

[0117] W为待查重功能点相似度;

[0118] X为待查重功能点级别相似度;

[0119] Y为待查重功能点词向量相似度;

[0120] Z为待查重功能点类型相似度。

[0121] 如图10所示,本申请另一方面还提供了一种基于神经网络的软件功能点查重装置,包括:

[0122] 功能点明细表获取模块,用于获取基于功能点法的软件功能点明细表,所述软件

功能点明细表包括功能点层级结构信息、功能点类别信息；

[0123] 历史功能点特征指标计算模块,用于利用软件功能点明细表对模型进行训练,得到训练好的模型和带有标注信息的功能点明细表,所述带有标注信息的功能点明细表作为历史功能点特征指标,所述历史功能点特征指标包括历史功能点级别、历史功能点词向量、历史功能点类型；

[0124] 待查重功能点特征指标计算模块,用于将待查重软件功能点明细表输入训练好的模型,输出待查重软件功能点明细表中每一个功能点的待查重功能点特征指标,所述待查重功能点特征指标包括待查重功能点级别、待查重功能点词向量、待查重功能点类型；

[0125] 相似度计算模块,用于计算所述历史功能点特征指标与所述待查重功能点特征指标的相似度；

[0126] 功能点查重结果输出模块,用于根据所述相似度和设定阈值确定待查重软件功能点明细表中的相似功能点,输出带有重复标注信息的软件功能点明细表,得到软件功能点查重结果。

[0127] 如图11所示,本申请的优选实施例还提供了一种电子设备,包括存储器、处理器及存储在存储器上并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上述实施例中的基于神经网络的软件功能点查重方法的步骤。

[0128] 如图12所示,本申请的优选实施例还提供了一种计算机设备,该计算机设备可以是终端或活体检测服务器,其内部结构图可以如图10所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。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和计算机程序。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。该计算机设备的网络接口用于与外部的其他计算机设备通过网络连接通信。该计算机程序被处理器执行时以实现上述基于神经网络的软件功能点查重方法的步骤。

[0129] 本领域技术人员可以理解,图12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。

[0130] 本申请的优选实施例还提供了一种存储介质,所述存储介质包括存储的程序,在所述程序运行时控制所述存储介质所在的设备执行上述实施例中的基于神经网络的软件功能点查重方法的步骤。

[0131] 需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0132] 本实施例方法所述功能若以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个或者多个计算设备可读取存储介质中。基于这样的理解,本申请实施例对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该软件产品存储在一个存储介质中,包括若干指令用以使得一台计算设备(可以是个人计算机,服务器,移动计算设备或者网络设备)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介

质。

[0133] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。本申请实施例中的方案可以采用各种计算机语言实现,例如,面向对象的程序设计语言Java和直译式脚本语言JavaScript等。

[0134] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0135] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0136] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0137] 尽管已描述了本申请的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例作出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0138] 显然,本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样,倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内,则本申请也意图包含这些改动和变型在内。

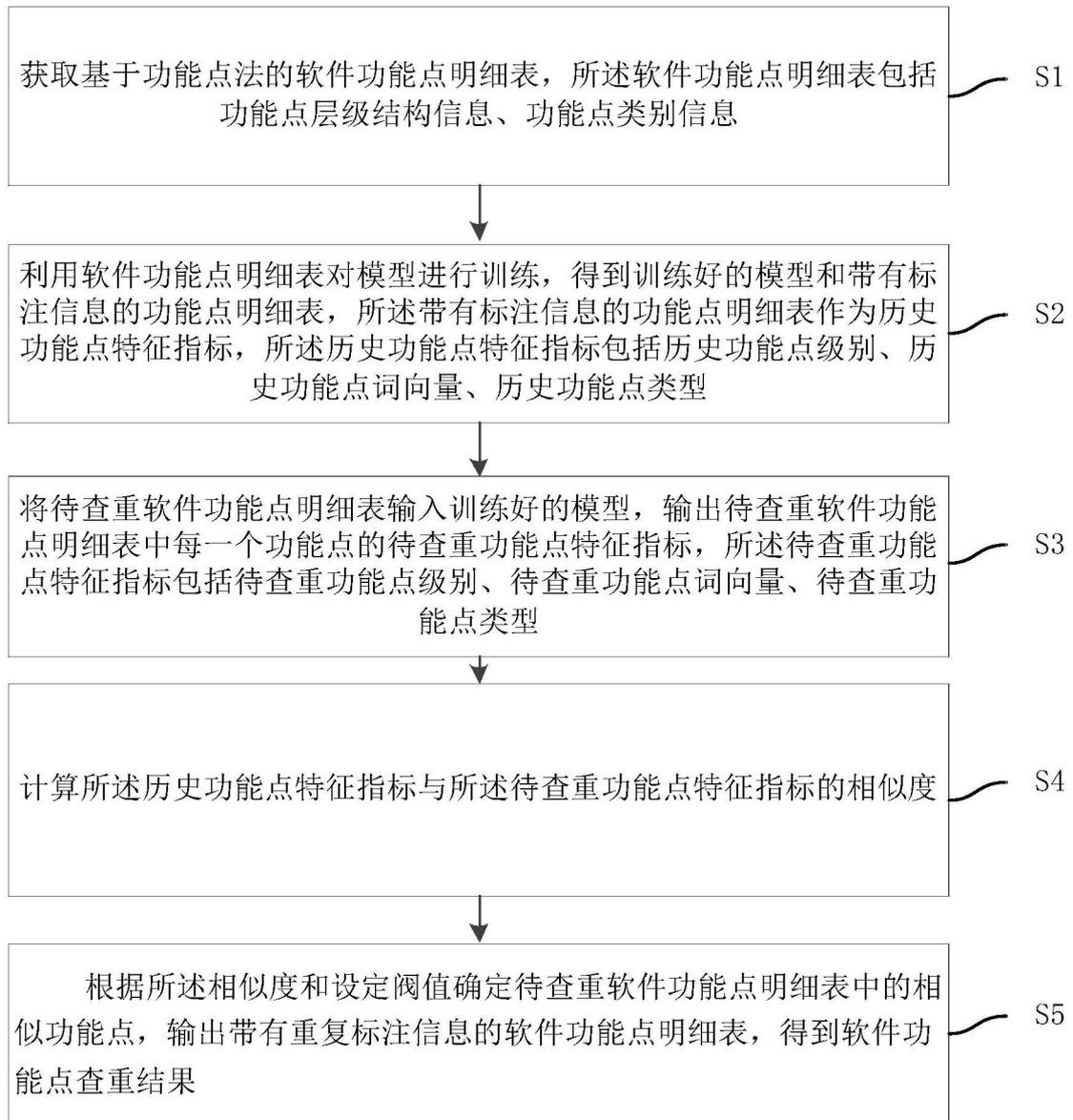


图1

软件开发工作量评估明细表 (功能点法)																		
序号	功能模块	一级功能	二级功能	三级功能	业务功能及功能点			原始评估										
					功能点名称	FTR/RET名称	DET字段	开发类型	类型	复杂度	FTR/RET	DET	UFP	折算后功能点	备注			
1	人脸基础信息管理	人脸本地管理	人脸基础管理	人脸基础管理	用户组创建	用户分组信息表	用户组ID、组名称、创建时间、创建者	新增	EI	低	1	4	3	3				
2					用户组删除	用户分组信息表	用户组ID、组名称、创建时间、创建者	新增	EI	低	1	4	3	3				
3					用户组修改	用户分组信息表	用户组ID、组名称、创建时间、创建者	新增	EI	低	1	4	3	3				
4					用户组查询	分组信息表	分组ID、分组名称、描述	新增	EQ	低	1	3	3	3				
16	认证中心	人脸认证服务	认证服务管理	认证服务管理	加密机人证认证接口配置文件	接口配置文件	接口APPID、接口APPKEY、接口地址、配置时间、启用开关	新增	UIF	低	1	5	7	7				
17					加密机认证结果获取入库	认证日志接口	流水号、用户ID、认证结果、操作时间、描述	新增	EI	低	1	5	3	3				
18					加密机认证数据签名查询			新增	EQ	低	1	5	3	3				
19					人脸识别信息查询	用户信息表、人脸基础数据表	用户ID、用户姓名、手机号、操作模块、操作结果、操作时间、终端标识、创建时间、创建人、修改时间、修改人、描述	新增	EQ	中	2	12	4	4				
20					用户信息缓存入库			新增	EI	低	1	5	3	3				
21					非正常状态用户过滤查询	登录日志表、用户信息表、用户标签表	用户ID、用户姓名、手机号、操作模块、操作结果、操作时间、终端标识、创建时间、创建人、修改时间、修改人、描述、标签	新增	EQ	中	3	13	4	4				
22					用户信息缓存更新			新增	EI	低	1	5	3	3				
23					用户信息来源标识查询	登录日志表、用户信息表、用户标签表	用户ID、用户姓名、手机号、操作模块、操作结果、操作时间、终端标识、创建时间、创建人、修改时间、修改人、描述、标签	新增	EQ	中	3	13	4	4				
24					用户信息缓存同步时间配置文件	接口配置文件	同步字段、接口APPKEY、接口地址、配置时间、启用开关	新增	UIF	低	1	5	7	7				
25					人脸识别解绑			新增	EQ	低	2	5	3	3				
26					认证日志记录	日志接口	用户ID、身份证号、手机号码、部门ID、登录时间、认证结果	新增	EI	低	1	6	3	3				
27					人脸登录日志上传	人脸登录日志表	认证方式、终端类型、用户ID、手机号码、登录结果、登录时间	新增	EI	低	1	6	3	3				
28					人脸登录日志报表配置表	登录日志表、用户信息表	用户ID、用户姓名、手机号、操作模块、操作结果、操作时间、终端标识、创建时间、创建人、修改时间、修改人、描述	新增	UIF	低	2	12	7	7				
29					人脸登录日志导出格式配置表	登录日志表、用户信息表	用户ID、用户姓名、手机号、操作模块、操作结果、操作时间、终端标识、创建时间、创建人、修改时间、修改人、描述	新增	UIF	低	2	12	7	7				
30					人脸基础数据格式定义配置表	数据格式定义配置表	APPID、APPKEY、数据格式、匹配正则	新增	UIF	低	1	4	7	7				
31					认证成功数据初始化			新增	EI	中	2	10	4	4				
32					认证成功数据下发			新增	EQ	低	1	5	3	3				
33					转储数据格式定义	转储数据表	用户ID、转储数据、转储格式、转储时间	新增	EI	低	1	4	3	3				
34					照片转储文件同步配置表	转储数据表	ID、端口、APPID、APPKEY	新增	UIF	低	1	4	7	7				
35					转储数据清理结果入库			新增	EI	低	1	5	3	3				
36					控制策略管理	DPS人脸认证规范存集群	DPS人脸认证规范存集群	DPS认证规范配置表	接口配置文件	接口APPID、接口APPKEY、接口地址、配置时间、启用开关	新增	UIF	低	1	5	7	7	
37								DPS数据同步查询			新增	EQ	低	1	5	3	3	
38								集群日志提交	系统日志表	集群ID、用户ID、身份证号、手机号码、同步时间、更新时间	新增	EI	低	1	6	3	3	
39								DPS认证规范配置表	接口配置文件	接口APPID、接口APPKEY、接口地址、配置时间、启用开关	新增	UIF	低	1	5	7	7	
40						集群数据同步修改			新增	EI	低	1	5	3	3			
41						集群日志提交	系统日志表	集群ID、用户ID、身份证号、手机号码、同步时间、更新时间	新增	EI	低	1	6	3	3			
42						DPS认证规范配置表	接口配置文件	接口APPID、接口APPKEY、接口地址、配置时间、启用开关	新增	UIF	低	1	5	7	7			
43						集群数据同步查询			新增	EQ	低	1	5	3	3			
44	认证服务规范存集群	认证服务规范存集群	认证服务规范存集群	集群ID、用户ID、身份证号、手机号码、同步时间、更新时间	新增	EI	低	1	6	3	3							
小计										197								
合计 (乘以调整因子)										213								

图2

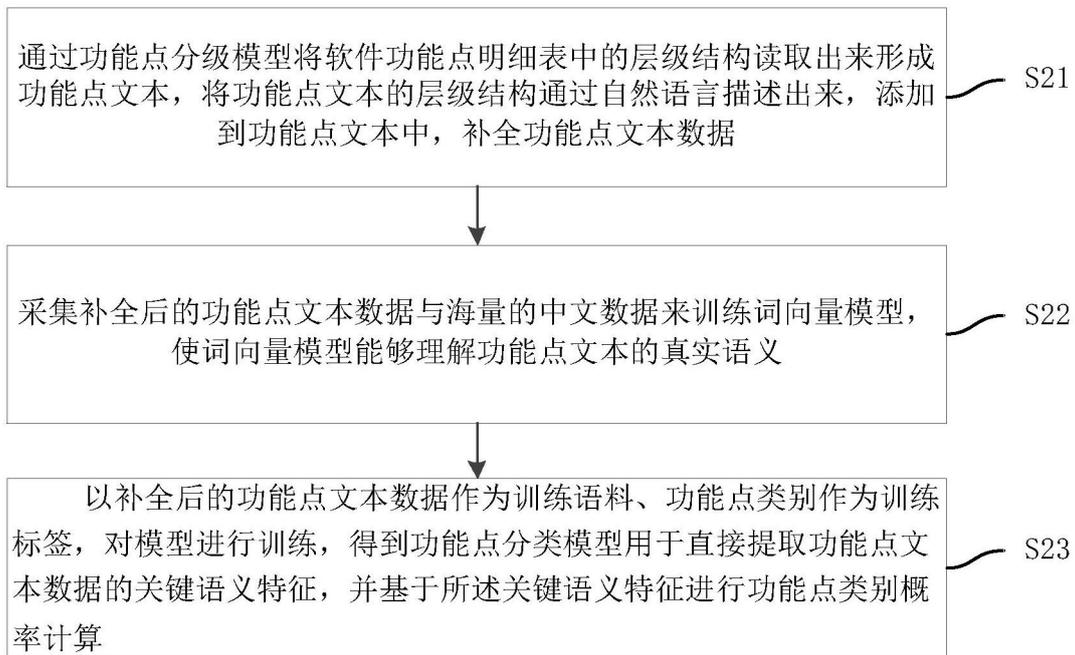


图3

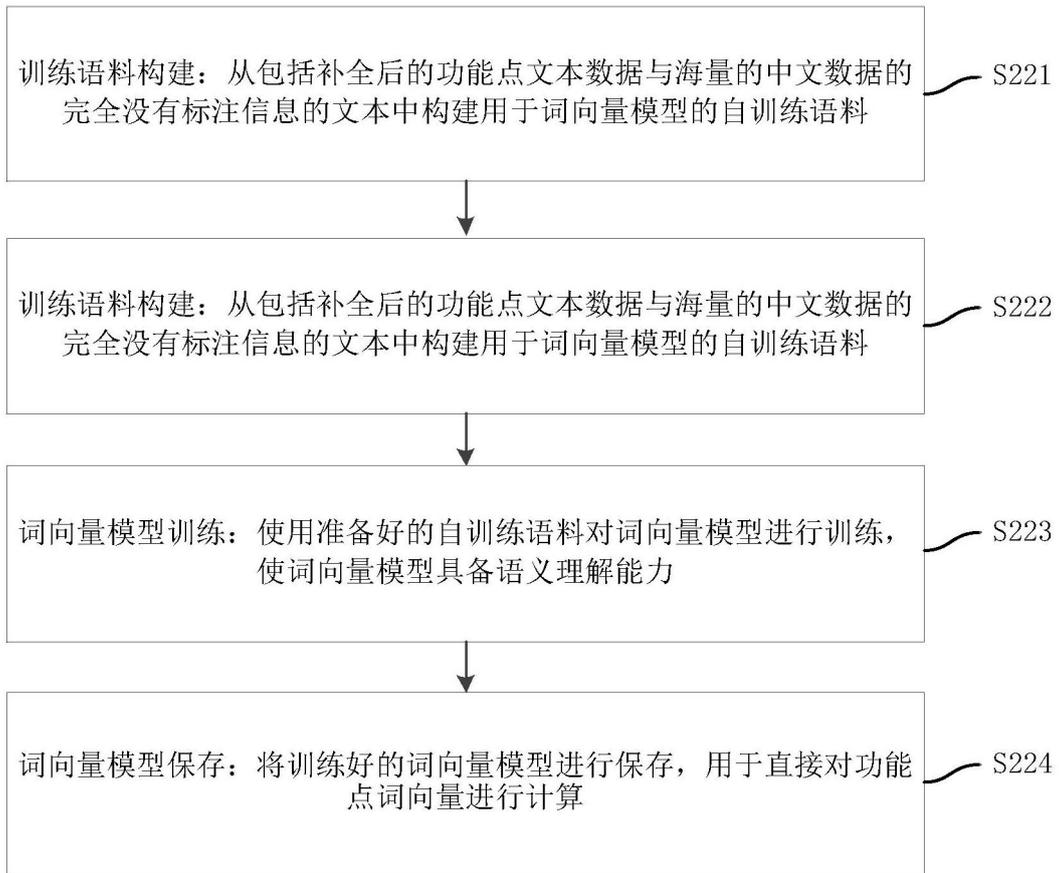


图4

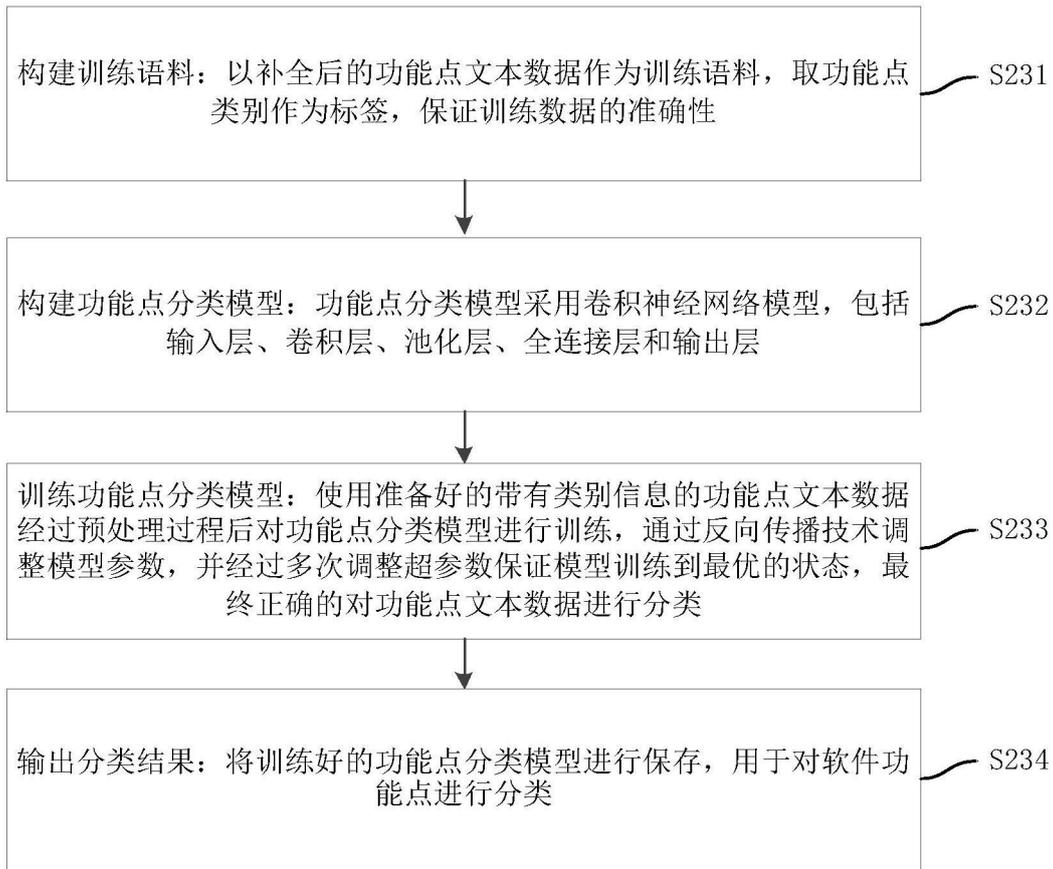


图5

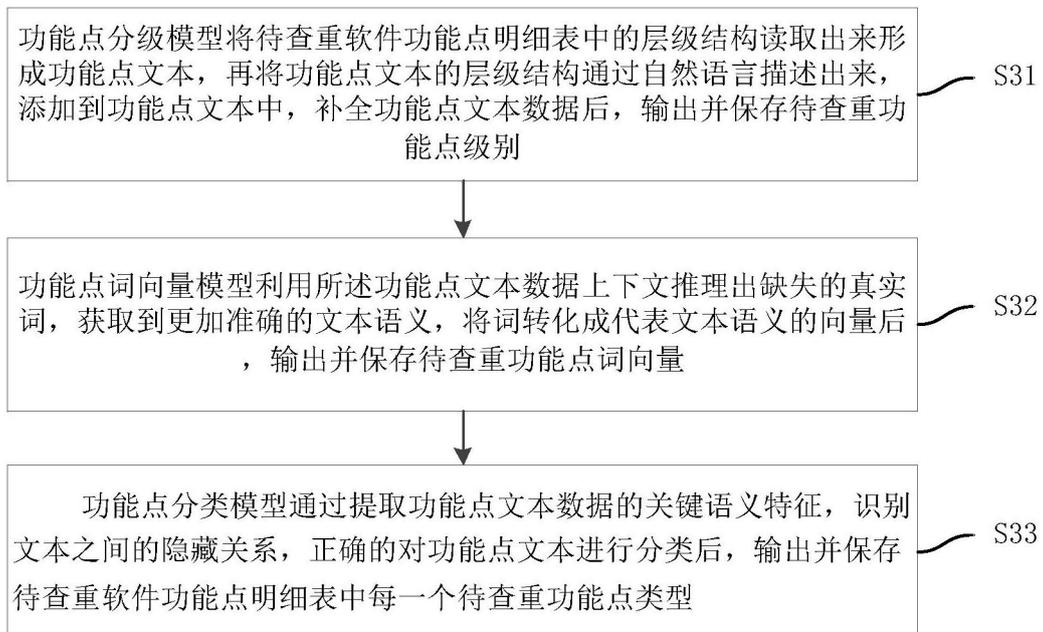


图6

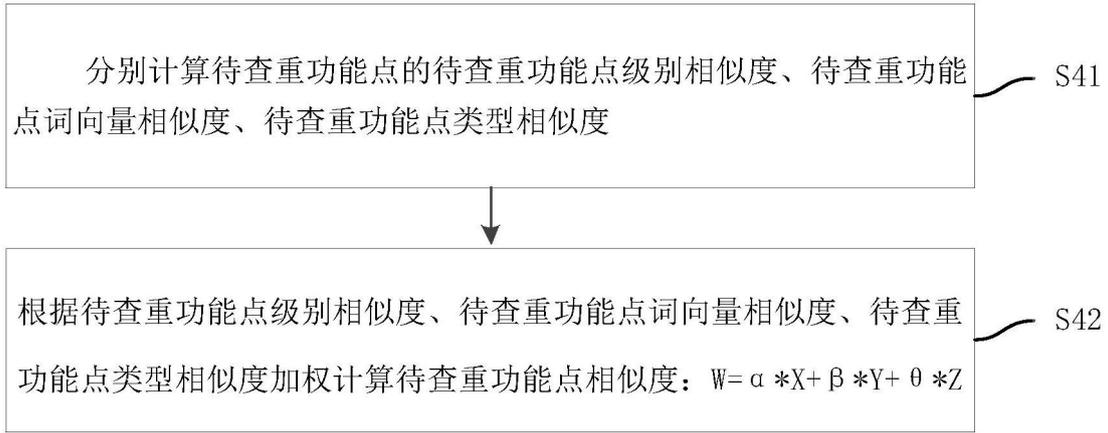


图7

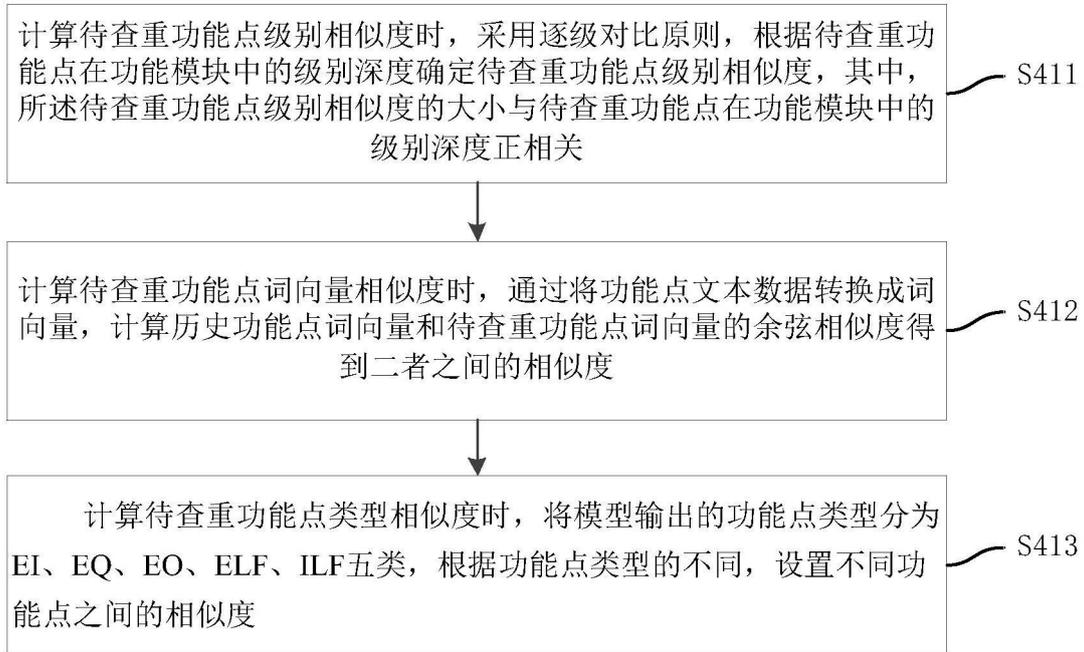


图8

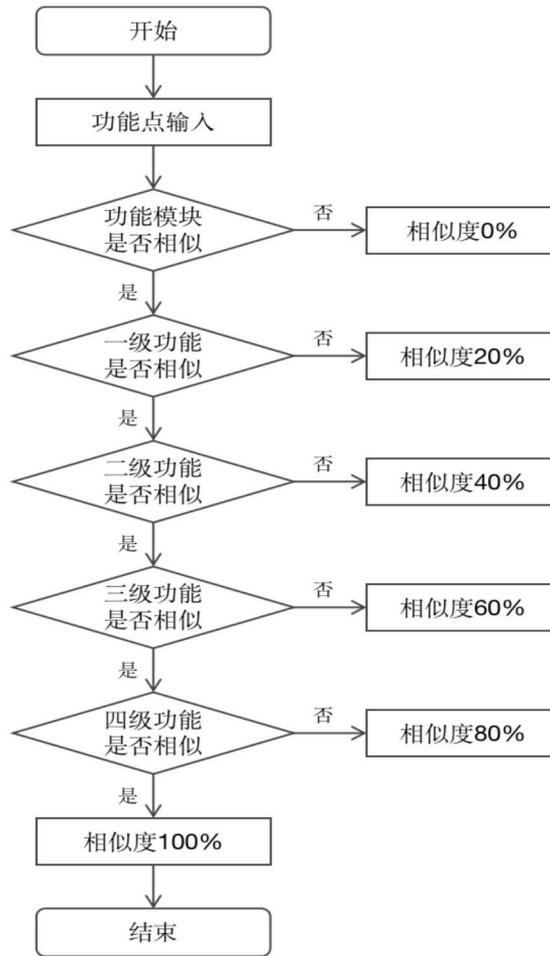


图9



图10



图11

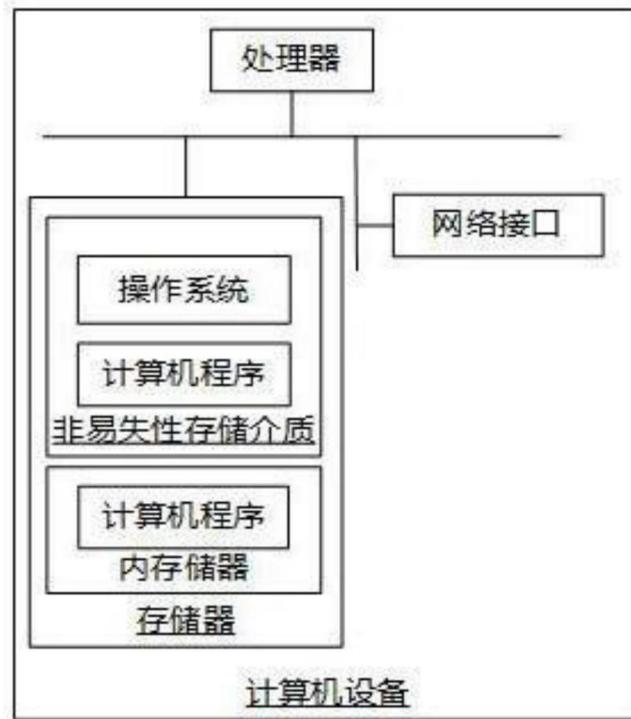


图12