



(12) 发明专利申请

(10) 申请公布号 CN 113657504 A

(43) 申请公布日 2021.11.16

(21) 申请号 202110946578.6

(22) 申请日 2021.08.18

(71) 申请人 上海东普信息科技有限公司

地址 201700 上海市青浦区外青松公路
5045号508室U区44号

(72) 发明人 徐梦佳 杨周龙 李斯

(51) Int.Cl.

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

权利要求书3页 说明书10页 附图5页

(54) 发明名称

图像检索方法、装置、计算机设备和存储介质

(57) 摘要

本发明涉及人工智能技术领域，公开了一种图像检索方法、装置、计算机设备和存储介质。该方法基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息；搭建并训练Faiss模型，优化所述Faiss模型，将所述对象特征向量与属性信息输入所述Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出。上述方法通过模型在图片库中找出匹配的无头件图片，得到对应的物品名称和编号等，核实之后到问题件仓库查找该物品，重新打包送至收件人，准确率高、速度快，减少了大量的人力成本。



1.一种图像检索方法,其特征在于,所述图像检索方法包括:

通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;

基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;

搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;

利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;

获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。

2.根据权利要求1所述的图像检索方法,其特征在于,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:

获取所述对象特征向量与属性信息;

将所述对象特征向量与属性信息压缩、编码、生成子聚类中心;

计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;

判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;

将所述查询索引与索引库中的本地索引集逐一遍历比对;

判断所述本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

3.根据权利要求1所述的图像检索方法,其特征在于,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:

获取所述对象特征向量与属性信息;

标记所述对象特征向量与属性信息,对所述对象特征向量压缩、编码、利用聚类算法生成样本类中心;

根据查询特征,匹配对应的样本类中心,计算残差特征;

根据所述残差特征,计算子聚类中心距离;

根据所述子聚类中心距离,判断素有子空间是否遍历完,若遍历完,则生成查询残差索引,若未遍历完,则继续遍历;

将所述查询残差索引与同类别的训练残差索引集逐一遍历比对;

判断所述同类别的训练残差索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

4.根据权利要求1所述的图像检索方法,其特征在于,所述基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息还包括:

对获取的图片进行预处理,筛选和提取出有效的数据样本;

对图片样本不平衡进行数据增强操作;

对ResNet50卷积神经网络进行调整,提取预训练所述ResNet50卷积神经网络的权重参

数作为初始值进行迁移学习。

5. 根据权利要求4所述的图像检索方法,其特征在于,所述对ResNet50卷积神经网络进行调整包括:

设置第一层卷积层中卷积核大小为 7×7 , stride为4,通道数为64,激活函数为Relu;

设置Max pool层中核大小为 3×3 , stride为2;

在残差模块stride=2的 1×1 的卷积核做下采样前加入一层 2×2 的平均池化层整合空间信息。

6. 根据权利要求4所述的图像检索方法,其特征在于,所述对ResNet50卷积神经网络进行调整包括:

在组块1的第一个bottleneck残差结构的输出添加64维的 1×1 卷积层跨越连接至组块4的第一个bottleneck残差结构;

在组块1的第二个bottleneck残差结构的输出添加128维的 1×1 卷积层跨越连接至组块4的第二个bottleneck残差结构。

7. 一种图像检索装置,其特征在于,所述图像检索装置包括:

获取模块,用于通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;

提取模块,用于基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;

搭建模块,用于搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;

检验模块,用于利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;

匹配模块,用于获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。

8. 根据权利要求7所述的图像检索装置,其特征在于,所述搭建模块,包括:

获取子模块,用于获取所述对象特征向量与属性信息;

生成子模块,用于将所述对象特征向量与属性信息压缩、编码、生成子聚类中心;

遍历子模块,用于计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;

判断子模块,用于判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;

比对子模块,用于将所述查询索引与索引库中的本地索引集逐一遍历比对;

输出子模块,用于判断所述本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

9. 一种计算机设备,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述计算机可读指令被所述处理器执行时,使得所述处理器执行如权利要求1至6中任一项所述的图像检索方法的步骤。

10. 一种存储有计算机可读指令的存储介质,所述计算机可读指令被一个或多个处理

器执行时,使得一个或多个处理器执行如权利要求1至6中任一项所述的图像检索方法的步骤。

图像检索方法、装置、计算机设备和存储介质

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及计算机视觉和深度学习领域的图像检索方法、装置、计算机设备和存储介质。

背景技术

[0002] 由于电商蓬勃发展,大家越来越习惯于网络购物或者邮寄商品,这就离不开快递业务。部分快递件在运输途中,可能出现贴到快递件上的快递运单磨损、脱落等情况,导致快递单号看不清,中转站或者派件公司将无法派送没有收件人信息的快递件。这时操作人员会实时拍照保存在无头件图片库,记录物品的详细信息,当有人认领的时候,再去相应的库中搜索。随着无头件的数据量越来越大,现有人工搜索效率不高且容易出错。

发明内容

[0003] 有鉴于此,有必要针对现有人工搜索效率不高且容易出错的技术问题,提供一种图像检索方法、装置、计算机设备和存储介质。

[0004] 一种图像检索方法,包括:通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。

[0005] 在其中一个实施例中,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:获取所述对象特征向量与属性信息;将所述对象特征向量与属性信息压缩、编码、生成子聚类中心;计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;将所述查询索引与索引库中的本地索引集逐一遍历比对;判断所述本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0006] 在其中一个实施例中,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:获取所述对象特征向量与属性信息;标记所述对象特征向量与属性信息,对所述对象特征向量压缩、编码、利用聚类算法生成样本类中心;根据查询特征,匹配对应的样本类中心,计算残差特征;根据所述残差特征,计算子聚类中心距离;根据所述子聚类中心距离,判断素有子空间是否遍历完,若遍历完,则生成查询残差索引,若未遍历完,则继续遍

历;将所述查询残差索引与同类别的训练残差索引集逐一遍历比对;判断所述同类别的训练残差索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0007] 在其中一个实施例中,所述基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息还包括:对获取的图片进行预处理,筛选和提取出有效的数据样本;对图片样本不平衡进行数据增强操作;对ResNet50卷积神经网络进行调整,提取预训练所述ResNet50卷积神经网络的权重参数作为初始值进行迁移学习。

[0008] 在其中一个实施例中,所述对ResNet50卷积神经网络进行调整包括:设置第一层卷积层中卷积核大小为 7×7 ,stride为4,通道数为64,激活函数为Relu;设置Max pool层中核大小为 3×3 ,stride为2;在残差模块stride=2的 1×1 的卷积核做下采样前加入一层 2×2 的平均池化层整合空间信息。

[0009] 在其中一个实施例中,所述对ResNet50卷积神经网络进行调整包括:在组块1的第一个bottleneck残差结构的输出添加64维的 1×1 卷积层跨越连接至组块4的第一个bottleneck残差结构;在组块1的第二个bottleneck残差结构的输出添加128维的 1×1 卷积层跨越连接至组块4的第二个bottleneck残差结构。

[0010] 一种图像检索装置,包括:获取模块,用于通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;提取模块,用于基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;检验模块,用于利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;匹配模块,用于获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。

[0011] 在其中一个实施例中,所述搭建模块,包括:获取子模块,用于获取所述对象特征向量与属性信息;生成子模块,用于将所述对象特征向量与属性信息压缩、编码、生成子聚类中心;遍历子模块,用于计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;判断子模块,用于判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;比对子模块,用于将所述查询索引与索引库中的本地索引集逐一遍历比对;输出子模块,用于判断所述本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0012] 一种计算机设备,包括存储器和处理器,所述存储器中存储有计算机可读指令,所述计算机可读指令被所述处理器执行时,使得所述处理器执行上述图像检索方法的步骤。

[0013] 一种存储有计算机可读指令的存储介质,所述计算机可读指令被一个或多个处理器执行时,使得一个或多个处理器执行上述图像检索方法的步骤。

[0014] 上述图像检索方法、装置、计算机设备和存储介质,通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计

算相似度并筛选出最相似的1至10张图片并输出；利用所述测试图片集检验训练的所述Faiss模型效果，若效果达到预定需求，则停止模型训练，得到度量模型；获取客户发送的一张图片，输入所述图片，应用所述度量模型进行检索，输出所述图片库中与所述检索图片最相似的1至10张图片，匹配得到对应的物品名称和编号。本发明采用上述技术方案后，利用以图搜图技术，训练优化Faiss模型，当搜索无头件时，直接输入需要查找的物体图片，通过模型在图片库中找出匹配的无头件图片，得到对应的物品名称和编号等，核实之后到问题件仓库查找该物品，重新打包送至收件人，整个流程准确率高、速度快，减少了大量的人力成本，同时更好的为广大消费者服务。

附图说明

[0015] 通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本发明的限制。

- [0016] 图1为本发明实施例中图像检索方法的一个实施例示意图；
- [0017] 图2为本发明实施例中图像检索方法的一个实施例示意图；
- [0018] 图3为本发明实施例中图像检索方法的一个实施例示意图；
- [0019] 图4为本发明实施例中图像检索装置的一个实施例示意图；
- [0020] 图5为本发明实施例中搭建模块的一个实施例示意图。

具体实施方式

[0021] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

[0022] 本技术领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“”和“该”也可包括复数形式。应该进一步理解的是，本发明的说明书中使用的措辞“包括”是指存在特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。

[0023] 如图1所示，在一个实施例中，提出了一种图像检索方法，该图像检索方法包括：

[0024] 步骤S101，通过扫描快递单获取图片，将图片收集形成图片库，将图片库中的图片分为训练图片集和测试图片集；

[0025] 收集现有无头件模块图片库的图片，将图片随机划分为训练图片集和测试图片集，建议测试图片集占比小。例如，将图像数据集按80%和20%的比例，依次分为训练图片集和测试图片集。

[0026] 步骤S102，基于ResNet50卷积神经网络提取训练图片集的对象特征向量与属性信息；

[0027] 基于深度学习的图像特征提取方法，即采用深度特征的图像特征提取算法。深度特征是指以卷积神经网络 (Convolutional Neural Network, CNN) 模型对原始图像数据进行特征提取所得到的特征向量。深度卷积网络模型具有较强的表达能力，并且免去了繁重复杂的手工定制特征流程。随着硬件性能逐渐提升、模型计算不断优化，对比采用特征提取

算法所生成的手工定制特征，深度特征在经过相同的索引检索算法处理后，在基于内容图像检索(CBIR)的性能指标精确度与召回率上表现优异。

[0028] 步骤S103，搭建并训练Faiss模型，优化Faiss模型，将对象特征向量与属性信息输入Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出；

[0029] Faiss是Facebook AI研究团队开源的针对聚类和相似性搜索库，它包含一种在任意大小的向量集合中搜索直到可能不适合在RAM中的新算法。它还包含用于评估和参数调整的支持代码。Faiss是用C++编写的，带有Python/numpy的完整封装，并使用GPU来获得更高的内存带宽和计算吞吐量。

[0030] Faiss的核心就是索引(index)概念，它封装了一组向量，并且可以选择是否进行预处理，帮忙高效的检索向量。Faiss有多种类型的索引，例如最简单的索引类型indexFlatL2，这就是基于L2距离(欧式距离)的暴力检索。不管建立什么类型的索引，所有的索引构建都需要知道操作数据的维度d，这里d=2048。对于大部分索引类型而言，在建立的时候都包含了训练阶段，基于训练集来分析向量的分布，但是L2这个索引可以跳过。当索引被建立和训练之后，可以调用add, search着两种方法。如果存在的向量太多，通过暴力搜索索引indexFlatL2搜索时间会变长，还可以利用一种加速搜索的方法indexIVFFlat(倒排文件)。起始就是使用k-means建立聚类中心，然后通过查询最近的聚类中心，然后比较聚类中所有向量得到相似的向量。也可以使用精确搜索，选择Faiss.indexFlatL2(欧式距离)和Faiss.indexFlatIP(内积)这两种索引类型。在本发明使用后者。将上一步骤得到的特征向量转化成矩阵，调用Faiss.IndexFlatIP()函数，输出相应的索引值和相似度，然后进行模型训练。

[0031] 步骤S104，利用测试图片集检验训练的Faiss模型效果，若效果达到预定需求，则停止模型训练，得到度量模型；

[0032] 应用优化的Faiss模型对测试集的图片进行无头件检索，输出现有图片库中与该图片最相似的10张图(这里0到10可以自行设置)，测试现有模型的效果，若效果不好则继续优化。

[0033] 步骤S105，获取客户发送的一张图片，输入图片，应用度量模型进行检索，输出图片库中与检索图片最相似的1至10张图片，匹配得到对应的物品名称和编号。

[0034] 当客户发过来一张图片时，输入这张图片，应用Faiss模型进行无头件检索，输出现有图片库中与该图片最相似的10张图，得到对应的物品名称和编号等，核实之后到问题件仓库查找该物品，重新打包送至收件人。

[0035] 在一个实施例中，如图2所示，搭建并训练Faiss模型，优化所述Faiss模型，将所述对象特征向量与属性信息输入所述Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出包括：

[0036] 步骤S201，获取对象特征向量与属性信息；

[0037] 以基于目标图像来提取目标对象的对象特征向量与属性信息。由于目标图像中包含目标对象，所以可以提取目标图像中目标对象的对象特征向量与属性信息，其中，对象特征向量可以为目标对象的特征向量，属性信息为目标图像中目标对象的其他相关信息，例如目标对象发货点、所在网点等信息。

[0038] 步骤S202，将对象特征向量与属性信息压缩、编码、生成子聚类中心；

[0039] PQ算法在算法的索引构建(数据训练)阶段,进行码本(聚类)训练,并对训练样本的特征数据集压缩、编码、生成索引。CBIR特征提取方法得到N个训练样本的特征向量集X。假设特征向量的维度为 Ω ,将向量维度 Ω 平分切割成P个子向量,对同属一个子维度空间的子向量组进行K-Means聚类生成子聚类中心(码本中心)。

[0040] 步骤S203,计算子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;

[0041] Faiss是Facebook开源的一个近似搜索的代码实现,其中有一些实现的不完善的地方,例如乘积量化算法(Product Quantization算法)在分段个数较多时,由于L1缓存大小限制,计算顺序会严重影响到搜索性能。此外,由于原始实现仅对分段数为4的倍数时做了循环展开,不具有普适性,本发明对此进行了优化。首先优化计算顺序,计算子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历。然后,优化循环展开,将原始只针对分段数为四的倍数的循环展开扩展到对所有分段数的实现,使得实现具有更好的适用性。

[0042] 步骤S204,判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;

[0043] 优化循环展开,将原始只针对分段数为四的倍数的循环展开扩展到对所有分段数的实现,使得实现具有更好的适用性。

[0044] 步骤S205,将查询索引与索引库中的本地索引集逐一遍历比对;

[0045] 步骤S206,判断本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0046] PQ算法在算法的索引构建(数据训练)阶段,进行码本(聚类)训练,并对训练样本的特征数据集压缩、编码、生成索引。CBIR特征提取方法得到N个训练样本的特征向量集X。假设特征向量的维度为 Ω ,将向量维度 Ω 平分切割成P个子向量,对同属一个子维度空间的子向量组进行K-Means聚类生成子聚类中心(码本中心)。设置子聚类中心数量为M,以1到M之间的十进制数作为聚类中心的编号。再者,对所有样本特征向量的每个子向量标记出对应子维度空间中最近的聚类中心编号。最后,以特征向量为单位进行压缩编码,依据每个子向量所标记出的对应最近的十进制数聚类中心ID编号,并以此进行压缩编码,生成索引。至此,完成了训练特征向量集的量化编码。在得到外部检索请求之后,对查询样本进行同样的特征提取得到查询特征向量。其次,依据之前训练特征向量处理的过程,将查询特征向量切分成同样数量的子段,并通过先前训练特征向量集生成的子聚类中心进行量化编码,生成查询索引;最后,对查询索引与索引库中的本地索引集逐一遍历比对,根据汉明距离计算最近距离,选出Top10个最近距离的索引库中的索引,并以此作为检索结果返回。

[0047] 在一个实施例中,如图3所示,搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:

[0048] 步骤S301,获取对象特征向量与属性信息;

[0049] 以基于目标图像来提取目标对象的对象特征向量与属性信息。由于目标图像中包含目标对象,所以可以提取目标图像中目标对象的对象特征向量与属性信息,其中,对象特征向量可以为目标对象的特征向量,属性信息为目标图像中目标对象的其他相关信息,例

如目标对象发货点、所在网点等信息。

[0050] 步骤S302,标记对象特征向量与属性信息,对对象特征向量压缩、编码、利用聚类算法生成样本类中心;

[0051] 步骤S303,根据查询特征,匹配对应的样本类中心,计算残差特征;

[0052] 与正排索引相反,倒排索引虽然也会对每个文件生成文档ID以及数据内容进行分词,但在得到关键词后,还会对关键词进行词条ID编号,并以此作为键值索引(Index Term)组成索引表(Lexicon)。倒排索引的重点是在每个索引中记录每个关键词出现过的文档ID集合和对应的命中地址等位置信息(Posting Information),除此之外也包括词序、词频等常规统计信息。在检索查询时,对输入的数据进行分词得到词条ID编号,并仅在对应词条ID的索引中进行匹配查询,以此实现非线性搜索查询,最后根据位置信息与统计信息以某种度量方式和排序算法给出对应的文档作为检索结果。可以基于键值来查询文档。

[0053] 步骤S304,根据残差特征,计算子聚类中心距离;

[0054] 步骤S305,根据子聚类中心距离,判断素有子空间是否遍历完,若遍历完,则生成查询残差索引,若未遍历完,则继续遍历;

[0055] 利用Faiss框架对改良PQ索引检索算法实现。采用改良PQ索引检索算法的CBIR需要做到在不重建索引数据库的情况下对特定敏感数据进行索引入库,或是对索引数据库的数据实时动态增删。为避免算法重新训练并构建索引库所造成的高昂运维成本与耗时,并且考虑到RAM读写存储与GPU计算加速的优化需求,本发明采用Faiss框架实现了改良PQ编码的CBIR大规模索引检索算法,满足索引库数据的动态增删、索引库的RAM存放扩展、索引检索的GPU优化计算这三个需求,以降低索引数据库的运维复杂性,提高算法效率。跟踪外部检索请求的数据流,可将本发明的基于Faiss框架的CBIR处理流程细分为如下步骤数据预处理、提取深度特征、压缩编码生成索引、索引入库与增删、检索并回调结果。

[0056] 步骤S306,将查询残差索引与同类别的训练残差索引集逐一遍历比对;

[0057] 深度残差网络ResNet允许网络可以尽可能的加深,但其将输入中的一部分数据不经过卷积网络,直接送入到输出,保留一部分原始信息,这种结构有效防止了反向传播时的梯度弥散问题,从而使得网络深度可以达到上百层甚至更深。ResNet网络能添加更深卷积层提升效果,其中的残差块结构起关键作用。

[0058] 步骤S307,判断同类别的训练残差索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0059] 利用倒排索引对PQ算法进行改进,应用倒排索引的改良乘积量化编码算法,有利于CBIR的大规模索引检索需求。在查询特征的索引与训练特征集的索引库建立之前,先通过对训练特征集粗聚类,即样本聚类,实现对感兴趣区域(ROI)的训练特征快速标记与区分。检索查询时,查询特征仅与相同粗聚类中心的训练特征集进行索引检索查询,以非线性检索的方式遍历索引库中的部分索引并计算距离与排序,以此达到降低计算量、减少耗时的目的。

[0060] 在一个实施例中,基于ResNet50卷积神经网络提取训练图片集的对象特征向量与属性信息还包括:

[0061] 步骤S401,对获取的图片进行预处理,筛选和提取出有效的数据样本;

[0062] 步骤S402,对图片样本不平衡进行数据增强操作;

- [0063] 对数据集进行预处理,对数据样本切割增强生成模型训练集。
- [0064] 步骤S403,对ResNet50卷积神经网络进行调整,提取预训练ResNet50卷积神经网络的权重参数作为初始值进行迁移学习。
- [0065] 在一个实施例中,对ResNet50卷积神经网络进行调整包括:
- [0066] 步骤S501,设置第一层卷积层中卷积核大小为 7×7 ,stride为4,通道数为64,激活函数为Relu;
- [0067] 步骤S502,设置Max pool层中核大小为 3×3 ,stride为2;
- [0068] ResNet网络第1层都是经过 10×10 卷积层,感受视野较大,对于ImageNet数据库中的图像提取特征足够,但在本发明中,快递单污损情况复杂且大多数污损非常小,为了更加精确的对污损点进行分类这就需要提取更多有效细微特征,所以本发明在设计网络时对第1层 10×10 卷积层进行了改进,使之更加适应污损点检测特征提取任务。本发明设计中用 7×7 的堆叠卷积层替换 10×10 卷积层,一方面卷积层使用了更多的非线性激活函数,使得判决函数更具有判决性;另一方面,还有效降低了计算参数量,假设卷积层的输入和输出的特征图大小相同为Z,一个 7×7 卷积层所以在不改变初始感受野的情况下改进网络第1层,能为污损检测带来更好的性能。在残差模块stride=2的 1×1 的卷积核做下采样前加入一层 2×2 的平均池化层整合空间信息。在污损检测问题中,污损占整幅图像面积较小,获取的信息占比小,为了避免无用信息冗余,本发明在残差模块stride=2的 1×1 的卷积核做下采样前又加入一层 2×2 的平均池化层整合空间信息,平均池化层没有参数,不改变全局参数量,同时可防止在该层出现过拟合,减少计算量的同时,提高检测速度。
- [0069] 在一个实施例中,对ResNet50卷积神经网络进行调整包括:
- [0070] 步骤S601,在组块1的第一个bottleneck残差结构的输出添加64维的 1×1 卷积层跨越连接至组块4的第一个bottleneck残差结构;
- [0071] 步骤S602,在组块1的第二个bottleneck残差结构的输出添加128维的 1×1 卷积层跨越连接至组块4的第二个bottleneck残差结构。
- [0072] Xception Module结合深度可分离卷积和基础残差模块Bottleneck残差结构,通过残差学习单元不断加深网络模型,提取了丰富的语义特征,在保证精度的情况下可以减少模型的参数,降低运算成本。同时,网络的多尺度融合有了更好的效果,经过模块的特征提取与融合后,高低分辨率的交互有了更好的结果输出。
- [0073] 在一个实施例中,如图4所示,提供了一种图像检索装置,该图像检索装置包括:获取模块,用于通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;提取模块,用于基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;检验模块,用于利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;匹配模块,用于获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。
- [0074] 在一个实施例中,如图5所示,所述搭建模块,包括:获取子模块,用于获取所述对象特征向量与属性信息;生成子模块,用于将所述对象特征向量与属性信息压缩、编码、生

成子聚类中心；遍历子模块，用于计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历；判断子模块，用于判断所有子空间是否遍历完，若遍历完，则生成查询索引，若未遍历完，则继续遍历；比对子模块，用于将所述查询索引与索引库中的本地索引集逐一遍历比对；输出子模块，用于判断所述本地索引集是否遍历完，若遍历完，则计算最近距离，输出1至10个最近距离的索引库中的索引，若未遍历完，则继续遍历。

[0075] 在一个实施例中，提出了一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行所述计算机可读指令时实现以下步骤：通过扫描快递单获取图片，将所述图片收集形成图片库，将所述图片库中的图片分为训练图片集和测试图片集；基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息；搭建并训练Faiss模型，优化所述Faiss模型，将所述对象特征向量与属性信息输入所述Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出；利用所述测试图片集检验训练的所述Faiss模型效果，若效果达到预定需求，则停止模型训练，得到度量模型；获取客户发送的一张图片，输入所述图片，应用所述度量模型进行检索，输出所述图片库中与所述检索图片最相似的1至10张图片，匹配得到对应的物品名称和编号。

[0076] 在一个实施例中，所述搭建并训练Faiss模型，优化所述Faiss模型，将所述对象特征向量与属性信息输入所述Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出包括：获取所述对象特征向量与属性信息；将所述对象特征向量与属性信息压缩、编码、生成子聚类中心；计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历；判断所有子空间是否遍历完，若遍历完，则生成查询索引，若未遍历完，则继续遍历；将所述查询索引与索引库中的本地索引集逐一遍历比对；判断所述本地索引集是否遍历完，若遍历完，则计算最近距离，输出1至10个最近距离的索引库中的索引，若未遍历完，则继续遍历。

[0077] 在一个实施例中，所述搭建并训练Faiss模型，优化所述Faiss模型，将所述对象特征向量与属性信息输入所述Faiss模型，计算相似度并筛选出最相似的1至10张图片并输出包括：获取所述对象特征向量与属性信息；标记所述对象特征向量与属性信息，对所述对象特征向量压缩、编码、利用聚类算法生成样本类中心；根据查询特征，匹配对应的样本类中心，计算残差特征；根据所述残差特征，计算子聚类中心距离；根据所述子聚类中心距离，判断素有子空间是否遍历完，若遍历完，则生成查询残差索引，若未遍历完，则继续遍历；将所述查询残差索引与同类别的训练残差索引集逐一遍历比对；判断所述同类别的训练残差索引集是否遍历完，若遍历完，则计算最近距离，输出1至10个最近距离的索引库中的索引，若未遍历完，则继续遍历。

[0078] 在一个实施例中，所述基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息还包括：对获取的图片进行预处理，筛选和提取出有效的数据样本；对图片样本不平衡进行数据增强操作；对ResNet50卷积神经网络进行调整，提取预训练所述ResNet50卷积神经网络的权重参数作为初始值进行迁移学习。

[0079] 在一个实施例中，所述对ResNet50卷积神经网络进行调整包括：设置第一层卷积层中卷积核大小为 7×7 , stride为4, 通道数为64, 激活函数为Relu；设置Max pool层中核大

小为 3×3 , stride为2;在残差模块stride=2的 1×1 的卷积核做下采样前加入一层 2×2 的平均池化层整合空间信息。

[0080] 在一个实施例中,所述对ResNet50卷积神经网络进行调整包括:在组块1的第一个bottleneck残差结构的输出添加64维的 1×1 卷积层跨越连接至组块4的第一个bottleneck残差结构;在组块1的第二个bottleneck残差结构的输出添加128维的 1×1 卷积层跨越连接至组块4的第二个bottleneck残差结构。

[0081] 在一个实施例中,提出了一种存储有计算机可读指令的存储介质,所述计算机可读指令被一个或多个处理器执行时,使得一个或多个处理器执行以下步骤:通过扫描快递单获取图片,将所述图片收集形成图片库,将所述图片库中的图片分为训练图片集和测试图片集;基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息;搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出;利用所述测试图片集检验训练的所述Faiss模型效果,若效果达到预定需求,则停止模型训练,得到度量模型;获取客户发送的一张图片,输入所述图片,应用所述度量模型进行检索,输出所述图片库中与所述检索图片最相似的1至10张图片,匹配得到对应的物品名称和编号。

[0082] 在一个实施例中,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:获取所述对象特征向量与属性信息;将所述对象特征向量与属性信息压缩、编码、生成子聚类中心;计算所述子聚类中心的距离并将训练图片按批、当前查询样本的每个子空间、查询样本的顺序进行遍历;判断所有子空间是否遍历完,若遍历完,则生成查询索引,若未遍历完,则继续遍历;将所述查询索引与索引库中的本地索引集逐一遍历比对;判断所述本地索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0083] 在一个实施例中,所述搭建并训练Faiss模型,优化所述Faiss模型,将所述对象特征向量与属性信息输入所述Faiss模型,计算相似度并筛选出最相似的1至10张图片并输出包括:获取所述对象特征向量与属性信息;标记所述对象特征向量与属性信息,对所述对象特征向量压缩、编码、利用聚类算法生成样本类中心;根据查询特征,匹配对应的样本类中心,计算残差特征;根据所述残差特征,计算子聚类中心距离;根据所述子聚类中心距离,判断素有子空间是否遍历完,若遍历完,则生成查询残差索引,若未遍历完,则继续遍历;将所述查询残差索引与同类别的训练残差索引集逐一遍历比对;判断所述同类别的训练残差索引集是否遍历完,若遍历完,则计算最近距离,输出1至10个最近距离的索引库中的索引,若未遍历完,则继续遍历。

[0084] 在一个实施例中,所述基于ResNet50卷积神经网络提取所述训练图片集的对象特征向量与属性信息还包括:对获取的图片进行预处理,筛选和提取出有效的数据样本;对图片样本不平衡进行数据增强操作;对ResNet50卷积神经网络进行调整,提取预训练所述ResNet50卷积神经网络的权重参数作为初始值进行迁移学习。

[0085] 在一个实施例中,所述对ResNet50卷积神经网络进行调整包括:设置第一层卷积层中卷积核大小为 7×7 , stride为4,通道数为64,激活函数为Relu;设置Max pool层中核大小为 3×3 , stride为2;在残差模块stride=2的 1×1 的卷积核做下采样前加入一层 2×2 的

平均池化层整合空间信息。

[0086] 在一个实施例中,所述对ResNet50卷积神经网络进行调整包括:在组块1的第1个bottleneck残差结构的输出添加64维的 1×1 卷积层跨越连接至组块4的第1个bottleneck残差结构;在组块1的第2个bottleneck残差结构的输出添加128维的 1×1 卷积层跨越连接至组块4的第2个bottleneck残差结构。

[0087] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:只读存储器(ROM, Read Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁盘或光盘等。

[0088] 以上所述实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0089] 以上所述实施例仅表达了本发明一些示例性实施例,其描述较为具体和详细,但并不能因此而理解为对本发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干变形和改进,这些都属于本发明的保护范围。因此,本发明专利的保护范围应以所附权利要求为准。



图1

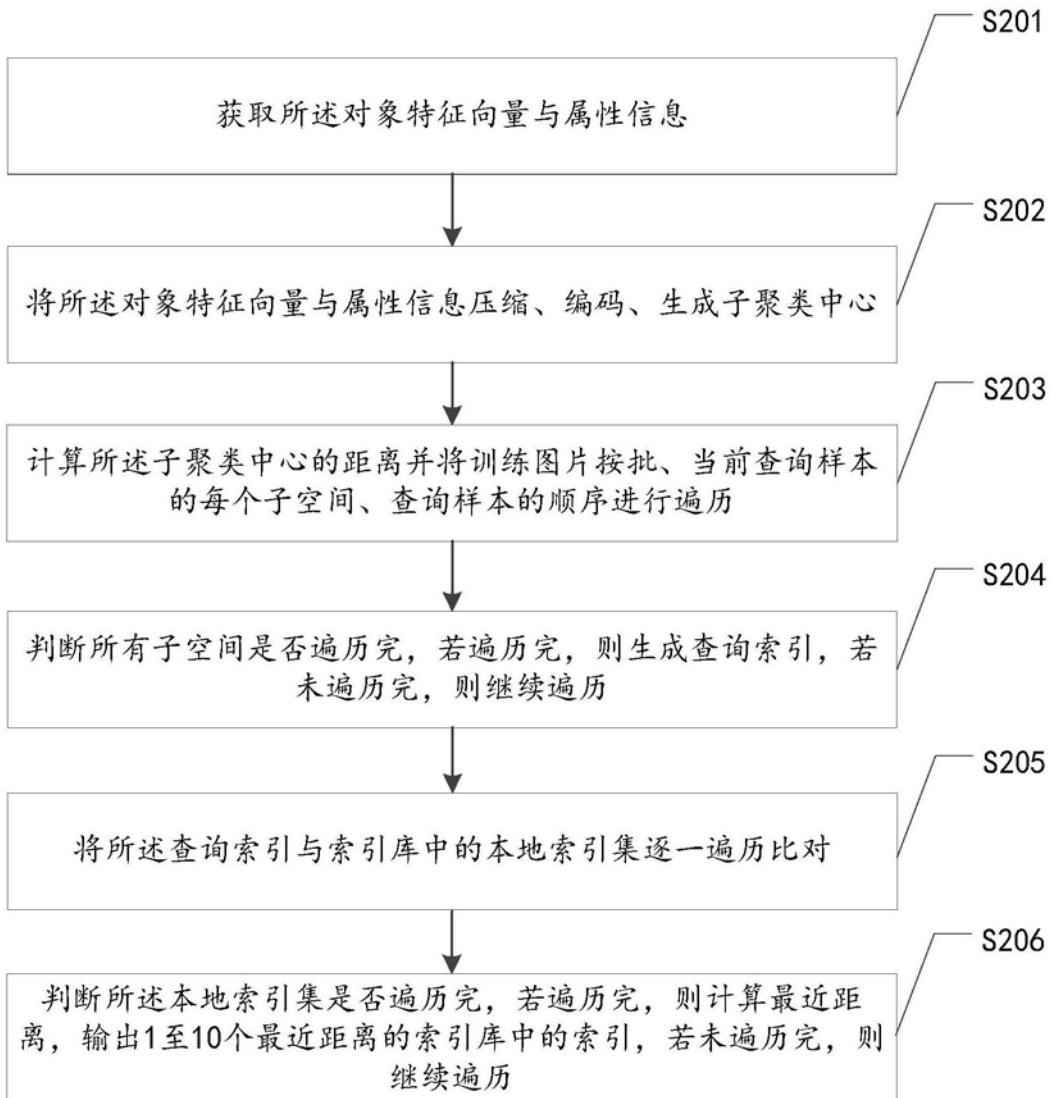


图2

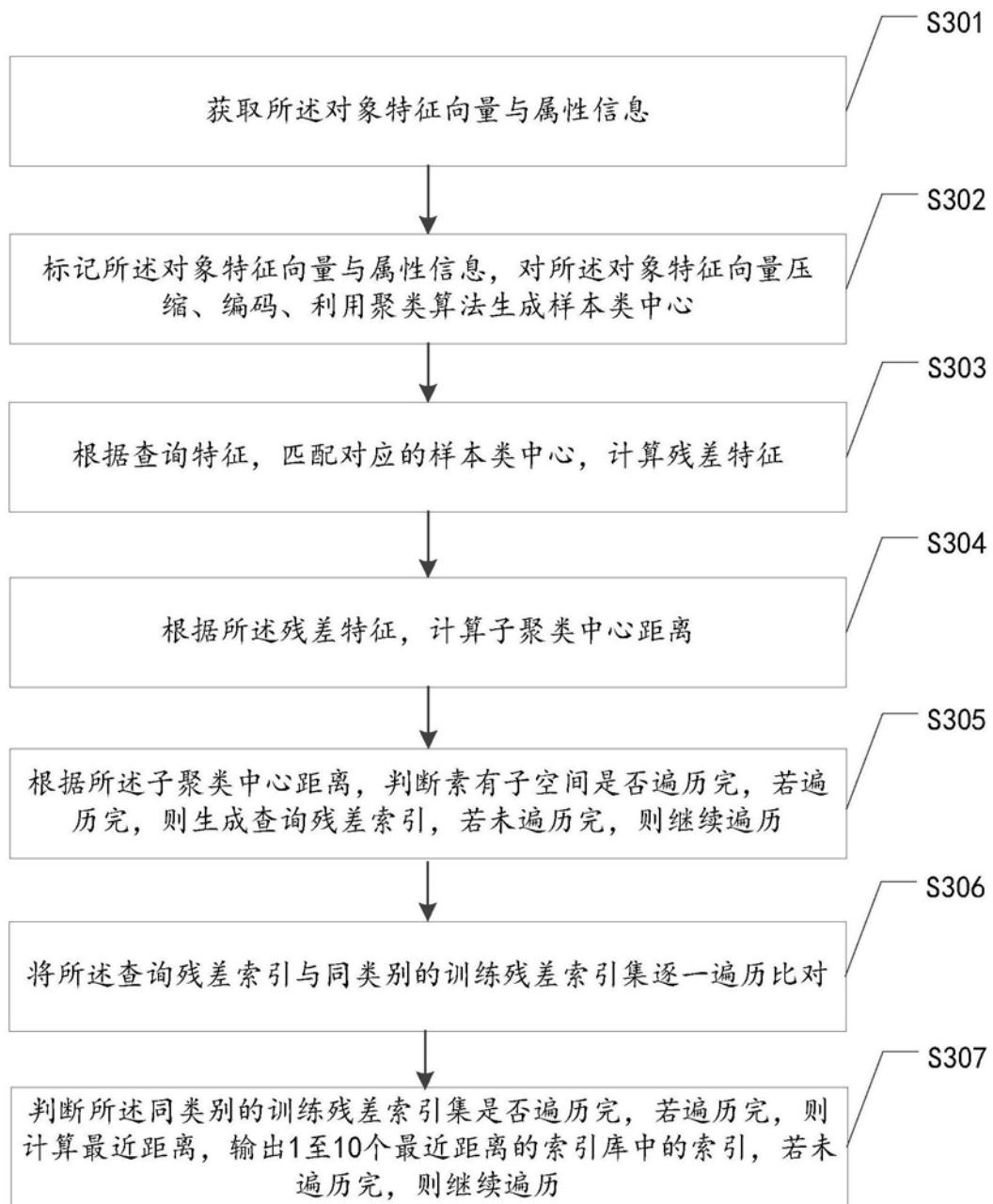


图3

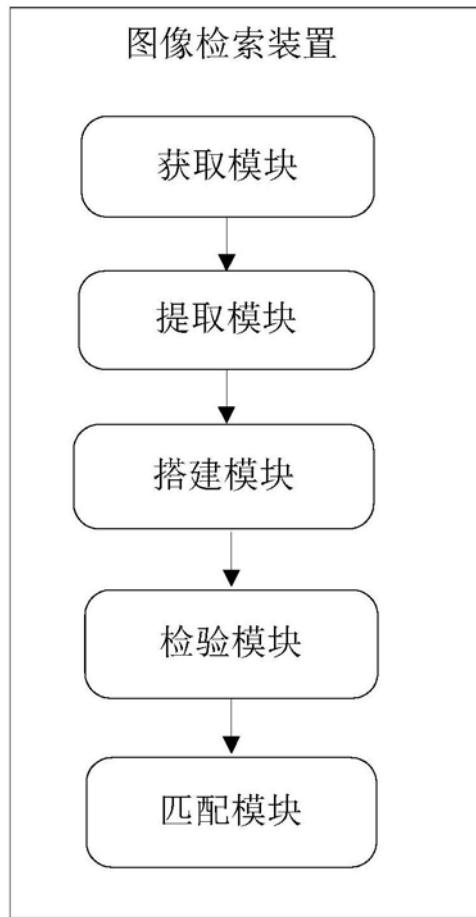


图4

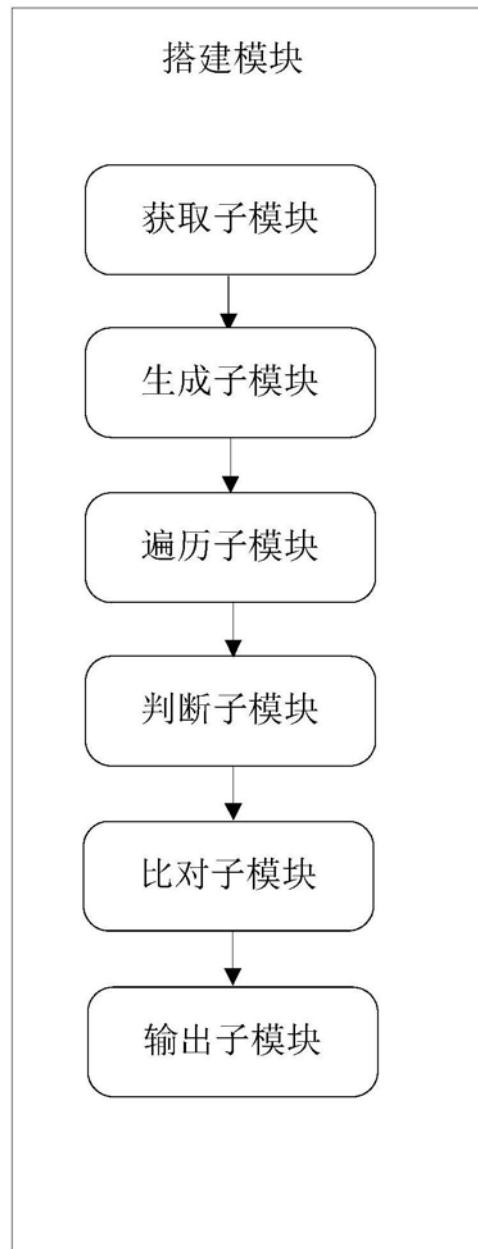


图5