



(12) 发明专利申请

(10) 申请公布号 CN 115795078 A

(43) 申请公布日 2023. 03. 14

(21) 申请号 202211572076.2

G06V 10/82 (2023.01)

(22) 申请日 2022.12.08

G06N 3/0464 (2022.01)

G06N 3/08 (2022.01)

(71) 申请人 北京沃东天骏信息技术有限公司

地址 100176 北京市北京经济技术开发区

科创十一街18号院2号楼4层A402室

申请人 北京京东世纪贸易有限公司

(72) 发明人 王倩芸 丁昆 刘朋樟 张屹峰

李阁 周梦迪 朱阳光 包勇军

(74) 专利代理机构 北京清亦华知识产权代理事

务所(普通合伙) 11201

专利代理师 石茵汀

(51) Int. Cl.

G06F 16/583 (2019.01)

G06V 10/44 (2022.01)

G06V 10/80 (2023.01)

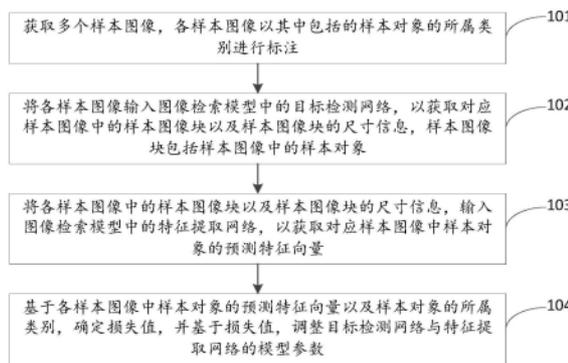
权利要求书5页 说明书18页 附图7页

(54) 发明名称

图像检索模型的训练方法、图像检索方法及装置

(57) 摘要

本申请提出一种图像检索模型的训练方法、图像检索方法及装置,涉及深度学习、图像处理等人工智能技术领域,图像检索模型的训练方法包括:获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注;将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息;将各样本图像块以及尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;基于各样本对象的预测特征向量以及所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。



1. 一种图像检索模型的训练方法,其特征在于,所述方法包括:

获取多个样本图像,各所述样本图像以其中包括的样本对象的所属类别进行标注;

将各所述样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及所述样本图像块的尺寸信息,所述样本图像块包括所述样本图像中的样本对象;

将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;

基于各所述样本图像中样本对象的预测特征向量以及所述样本对象的所属类别,确定损失值,并基于所述损失值,调整所述目标检测网络与所述特征提取网络的模型参数。

2. 根据权利要求1所述的方法,其特征在于,所述将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量,包括:

对于每个所述样本图像,将其中的样本图像块的尺寸信息,输入所述特征提取网络包括的第一尺寸特征提取模块,以基于所述尺寸信息获取所述样本图像块的第一尺寸特征;

将所述样本图像块输入所述特征提取网络包括的图像特征提取模块,以对所述样本图像块进行图像特征提取,得到所述样本图像块的图像特征向量;

将所述第一尺寸特征和所述图像特征向量,输入所述特征提取网络中与所述第一尺寸特征提取模块和所述图像特征提取模块连接的第一融合模块,以将所述第一尺寸特征和所述图像特征向量融合,得到所述样本图像中样本对象的预测特征向量。

3. 根据权利要求2所述的方法,其特征在于,所述图像特征提取模块包括依次连接的图像处理子模块以及第一主干网络;

所述将所述样本图像块输入所述特征提取网络包括的图像特征提取模块,以对所述样本图像块进行图像特征提取,得到所述样本图像块的图像特征向量,包括:

将所述样本图像块输入所述图像处理子模块,以对所述样本图像块进行尺寸缩放,得到预设尺寸的第一图像块,并获取所述第一图像块对应的第一张量;

将所述第一张量输入所述第一主干网络,以基于所述第一张量进行图像特征提取,得到所述样本图像块的图像特征向量。

4. 根据权利要求2所述的方法,其特征在于,所述第一融合模块包括依次连接的多层感知机子模块以及融合子模块;

所述将所述第一尺寸特征和所述图像特征向量,输入所述特征提取网络中与所述第一尺寸特征提取模块和所述图像特征提取模块连接的第一融合模块,以将所述第一尺寸特征和所述图像特征向量融合,得到所述样本图像中样本对象的预测特征向量,包括:

将所述第一尺寸特征输入所述多层感知机子模块,以对所述第一尺寸特征进行维度扩展,得到第二尺寸特征;

将所述第二尺寸特征和所述图像特征向量输入所述融合子模块,以将所述第二尺寸特征和所述图像特征向量融合,得到所述样本图像中样本对象的预测特征向量。

5. 根据权利要求2-4中任一项所述的方法,其特征在于,所述尺寸信息包括所述样本图像块的宽度和高度;所述基于所述尺寸信息获取所述样本图像块的第一尺寸特征,包括:

将所述样本图像块的宽度和高度,确定为所述第一尺寸特征;

或者,将所述样本图像块的宽度和高度的比值,确定为所述第一尺寸特征;

或者,确定各所述样本图像中样本图像块的宽度和高度的比值,并确定各所述比值的均值和标准差,基于所述比值、所述均值和所述标准差,确定所述第一尺寸特征;

或者,确定各所述样本图像中样本图像块的宽度和高度的对数比值,并确定各所述对数比值的均值和标准差,基于所述对数比值、各所述对数比值的均值和标准差,确定所述第一尺寸特征。

6. 根据权利要求1所述的方法,其特征在于,所述将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量,包括:

对于每个所述样本图像,将所述样本图像中样本图像块的尺寸信息输入所述特征提取网络包括的第二尺寸特征提取模块,以基于所述尺寸信息获取所述样本图像块的第三尺寸特征;

将所述样本图像块输入所述特征提取网络包括的图像处理模块,以对所述样本图像块进行尺寸缩放,得到预设尺寸的第二图像块,并获取所述第二图像块对应的第二张量;

将所述第三尺寸特征和所述第二张量,输入所述特征提取网络中与所述第二尺寸特征提取模块及所述图像处理模块连接的第二融合模块,以将所述第三尺寸特征和所述第二张量融合,得到融合特征向量;

将所述融合特征向量输入所述特征提取网络中与所述第二融合模块连接的第二主干网络,以基于所述融合特征向量进行特征提取,得到所述样本图像中样本对象的预测特征向量。

7. 根据权利要求6所述的方法,其特征在于,所述第二张量包括所述样本图像块中各像素点在多个通道的像素值;

所述将所述第三尺寸特征和所述第二张量,输入所述特征提取网络中与所述第二尺寸特征提取模块及所述图像处理模块连接的第二融合模块,以将所述第三尺寸特征和所述第二张量融合,得到融合特征向量,包括:

将所述第三尺寸特征和所述第二张量,输入所述第二融合模块,以对于所述样本图像块中的每个像素点,将所述第三尺寸特征与所述像素点在所述多个通道的像素值拼接,得到所述像素点在多个所述通道的特征向量;

基于所述样本图像块中的各个像素点在多个所述通道的特征向量,生成所述融合特征向量。

8. 一种图像检索方法,其特征在于,所述方法包括:

获取待检索的检索图像;

将所述检索图像输入图像检索模型中的目标检测网络,以获取所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,所述目标图像块包括所述检索图像中的目标对象;

将所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取所述目标对象的预测特征向量,其中,所述图像检索模型基于权利要求1-7中任一项所述的方法训练得到;

基于所述目标对象的预测特征向量,从多个候选图像中确定目标图像。

9. 根据权利要求8所述的方法,其特征在于,所述基于所述目标对象的预测特征向量,从多个候选图像中确定目标图像,包括:

获取各所述候选图像中候选对象的候选特征向量;

确定所述目标对象的预测特征向量与各所述候选图像中候选对象的候选特征向量之间的相似度;

基于所述目标对象的预测特征向量与各所述候选图像中候选对象的候选特征向量之间的相似度,从所述多个候选图像中确定所述目标图像。

10. 根据权利要求9所述的方法,其特征在于,所述获取各所述候选图像中候选对象的候选特征向量之前,还包括:

对于每个所述候选图像,将所述候选图像输入所述目标检测网络,以获取所述候选图像中的候选图像块以及所述候选图像块的尺寸信息,所述候选图像块包括所述候选图像中的候选对象;

将所述候选图像中的候选图像块以及所述候选图像块的尺寸信息,输入所述特征提取网络,以获取所述候选对象的预测特征向量。

11. 一种图像检索模型的训练装置,其特征在于,所述装置包括:

第一获取模块,用于获取多个样本图像,各所述样本图像以其中包括的样本对象的所属类别进行标注;

第一处理模块,用于将各所述样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及所述样本图像块的尺寸信息,所述样本图像块包括所述样本图像中的样本对象;

第二处理模块,用于将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;

模型参数调整模块,用于基于各所述样本图像中样本对象的预测特征向量以及所述样本对象的所属类别,确定损失值,并基于所述损失值,调整所述目标检测网络与所述特征提取网络的模型参数。

12. 根据权利要求11所述的装置,其特征在于,所述第二处理模块,包括:

第一尺寸特征提取单元,用于对于每个所述样本图像,将其中的样本图像块的尺寸信息,输入所述特征提取网络包括的第一尺寸特征提取模块,以基于所述尺寸信息获取所述样本图像块的第一尺寸特征;

图像特征提取单元,用于将所述样本图像块输入所述特征提取网络包括的图像特征提取模块,以对所述样本图像块进行图像特征提取,得到所述样本图像块的图像特征向量;

第一融合单元,用于将所述第一尺寸特征和所述图像特征向量,输入所述特征提取网络中与所述第一尺寸特征提取模块和所述图像特征提取模块连接的第一融合模块,以将所述第一尺寸特征和所述图像特征向量融合,得到所述样本图像中样本对象的预测特征向量。

13. 根据权利要求12所述的装置,其特征在于,所述图像特征提取模块包括依次连接的图像处理子模块以及第一主干网络;

所述图像特征提取单元,用于:

将所述样本图像块输入所述图像处理子模块,以对所述样本图像块进行尺寸缩放,得到预设尺寸的第一图像块,并获取所述第一图像块对应的第一张量;

将所述第一张量输入所述第一主干网络,以基于所述第一张量进行图像特征提取,得到所述样本图像块的图像特征向量。

14. 根据权利要求12所述的装置,其特征在于,所述第一融合模块包括依次连接的多层感知机子模块以及融合子模块;

所述第一融合单元,用于:

将所述第一尺寸特征输入所述多层感知机子模块,以对所述第一尺寸特征进行维度扩展,得到第二尺寸特征;

将所述第二尺寸特征和所述图像特征向量输入所述融合子模块,以将所述第二尺寸特征和所述图像特征向量融合,得到所述样本图像中样本对象的预测特征向量。

15. 根据权利要求12-14中任一项所述的装置,其特征在于,所述尺寸信息包括所述样本图像块的宽度和高度;所述第一尺寸特征提取单元,用于:

将所述样本图像块的宽度和高度,确定为所述第一尺寸特征;

或者,将所述样本图像块的宽度和高度的比值,确定为所述第一尺寸特征;

或者,确定各所述样本图像中样本图像块的宽度和高度的比值,并确定各所述比值的均值和标准差,基于所述比值、所述均值和所述标准差,确定所述第一尺寸特征;

或者,确定各所述样本图像中样本图像块的宽度和高度的对数比值,并确定各所述对数比值的均值和标准差,基于所述对数比值、各所述对数比值的均值和标准差,确定所述第一尺寸特征。

16. 根据权利要求11所述的装置,其特征在于,所述第二处理模块,包括:

第二尺寸特征提取单元,用于对于每个所述样本图像,将所述样本图像中样本图像块的尺寸信息输入所述特征提取网络包括的第二尺寸特征提取模块,以基于所述尺寸信息获取所述样本图像块的第三尺寸特征;

图像处理单元,用于将所述样本图像块输入所述特征提取网络包括的图像处理模块,以对所述样本图像块进行尺寸缩放,得到预设尺寸的第二图像块,并获取所述第二图像块对应的第二张量;

第二融合单元,用于将所述第三尺寸特征和所述第二张量,输入所述特征提取网络中与所述第二尺寸特征提取模块及所述图像处理模块连接的第二融合模块,以将所述第三尺寸特征和所述第二张量融合,得到融合特征向量;

融合特征提取单元,用于将所述融合特征向量输入所述特征提取网络中与所述第二融合模块连接的第二主干网络,以基于所述融合特征向量进行特征提取,得到所述样本图像中样本对象的预测特征向量。

17. 根据权利要求16所述的装置,其特征在于,所述第二张量包括所述样本图像块中各像素点在多个通道的像素值;

所述第二融合单元,用于:

将所述第三尺寸特征和所述第二张量,输入所述第二融合模块,以对于所述样本图像块中的每个像素点,将所述第三尺寸特征与所述像素点在所述多个通道的像素值拼接,得到所述像素点在多个所述通道的特征向量;

基于所述样本图像块中的各个像素点在多个所述通道的特征向量,生成所述融合特征向量。

18. 一种图像检索装置,其特征在於,所述装置包括:

第二获取模块,用于获取待检索的检索图像;

第三处理模块,用于将所述检索图像输入图像检索模型中的目标检测网络,以获取所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,所述目标图像块包括所述检索图像中的目标对象;

第四处理模块,用于将所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取所述目标对象的预测特征向量,其中,所述图像检索模型基于权利要求1-7中任一项所述的方法训练得到;

确定模块,用于基于所述目标对象的预测特征向量,从多个候选图像中确定目标图像。

19. 根据权利要求18所述的装置,其特征在於,所述确定模块,包括:

获取单元,用于获取各所述候选图像中候选对象的候选特征向量;

第一确定单元,用于确定所述目标对象的预测特征向量与各所述候选图像中候选对象的候选特征向量之间的相似度;

第二确定单元,用于基于所述目标对象的预测特征向量与各所述候选图像中候选对象的候选特征向量之间的相似度,从所述多个候选图像中确定所述目标图像。

20. 根据权利要求19所述的装置,其特征在於,还包括:

第三获取模块,用于对于每个所述候选图像,将所述候选图像输入所述目标检测网络,以获取所述候选图像中的候选图像块以及所述候选图像块的尺寸信息,所述候选图像块包括所述候选图像中的候选对象;

第五处理模块,用于将所述候选图像中的候选图像块以及所述候选图像块的尺寸信息,输入所述特征提取网络,以获取所述候选对象的预测特征向量。

21. 一种电子设备,其特征在於,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-7中任一项所述的方法,或者执行权利要求8-10中任一项所述的方法。

22. 一种存储有计算机指令的非瞬时计算机可读存储介质,其特征在於,所述计算机指令用于使所述计算机执行权利要求1-7中任一项所述的方法,或者执行权利要求8-10中任一项所述的方法。

## 图像检索模型的训练方法、图像检索方法及装置

### 技术领域

[0001] 本申请涉及深度学习、图像处理等人工智能技术领域,尤其涉及一种图像检索模型的训练方法、图像检索方法及装置。

### 背景技术

[0002] 图像检索技术在多种领域具有广泛的应用。比如在电商领域的同款推荐、属性治理、类目合并等场景中,需要通过图像检索技术,检索与已知图像相似的图像。

[0003] 相关技术中的图像检索技术,在模型训练阶段一般是将图像中的ROI(region of interest,感兴趣区域)缩放至固定尺寸后直接输入网络。由于图像的ROI的尺寸多样,其中ROI的尺寸反映了图像中主体对象的某些重要属性,比如ROI的孔径比(即宽度和高度的比例)反映了衣服的长短款,而将图像的ROI缩放至固定尺寸,会引起图像中主体对象的形变,从而上述将图像中的ROI缩放至固定尺寸后直接输入网络进行模型训练的方式,会导致模型无法学习到图像中主体对象的真实形态,在进行检索时容易检索到与已知图像不相关的图像,检索结果的准确性差。

### 发明内容

[0004] 本申请旨在至少在一定程度上解决相关技术中的技术问题之一。

[0005] 本申请提出一种图像检索模型的训练方法、图像检索方法及装置,以解决相关技术中训练得到的图像检索模型,在进行检索时容易检索到与已知图像不相关的图像,检索结果的准确性差的技术问题。

[0006] 本申请第一方面实施例提出了一种图像检索模型的训练方法,包括:获取多个样本图像,各所述样本图像以其中包括的样本对象的所属类别进行标注;将各所述样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及所述样本图像块的尺寸信息,所述样本图像块包括所述样本图像中的样本对象;将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;基于各所述样本图像中样本对象的预测特征向量以及所述样本对象的所属类别,确定损失值,并基于所述损失值,调整所述目标检测网络与所述特征提取网络的模型参数。由此,能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0007] 本申请第二方面实施例提出了一种图像检索方法,包括:获取待检索的检索图像;将所述检索图像输入图像检索模型中的目标检测网络,以获取所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,所述目标图像块包括所述检索图像中的目标对象;将所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取所述目标对象的预测特征向量,其中,所述图像检索模型基于第一方面实施例所述的方法训练得到;基于所述目标对象的预测特征向量,从多个候选图像中确定目标图像。由此,提高了图像检索的准确性。

[0008] 本申请第三方面实施例提出了一种图像检索模型的训练装置,包括:第一获取模块,用于获取多个样本图像,各所述样本图像以其中包括的样本对象的所属类别进行标注;第一处理模块,用于将各所述样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及所述样本图像块的尺寸信息,所述样本图像块包括所述样本图像中的样本对象;第二处理模块,用于将各所述样本图像中的样本图像块以及所述样本图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;模型参数调整模块,用于基于各所述样本图像中样本对象的预测特征向量以及所述样本对象的所属类别,确定损失值,并基于所述损失值,调整所述目标检测网络与所述特征提取网络的模型参数。

[0009] 本申请第四方面实施例提出了一种图像检索装置,包括:第二获取模块,用于获取待检索的检索图像;第三处理模块,用于将所述检索图像输入图像检索模型中的目标检测网络,以获取所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,所述目标图像块包括所述检索图像中的目标对象;第四处理模块,用于将所述检索图像中的目标图像块以及所述目标图像块的尺寸信息,输入所述图像检索模型中的特征提取网络,以获取所述目标对象的预测特征向量,其中,所述图像检索模型基于第一方面实施例所述的方法训练得到;确定模块,用于基于所述目标对象的预测特征向量,从多个候选图像中确定目标图像。

[0010] 本申请第五方面实施例提出了一种电子设备,包括:至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如本申请第一方面实施例提出的图像检索模型的训练方法,或者执行如本申请第二方面实施例提出的图像检索方法。

[0011] 本申请第六方面实施例提出了一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行如本申请第一方面实施例提出的图像检索模型的训练方法,或者执行如本申请第二方面实施例提出的图像检索方法。

[0012] 本申请第七方面实施例提出了一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现如本申请第一方面实施例提出的图像检索模型的训练方法本申请第二方面实施例提出的图像检索方法,或者执行如本申请第二方面实施例提出的图像检索方法。

[0013] 本申请附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本申请的实践了解到。

## 附图说明

[0014] 本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0015] 图1为本申请实施例一所提供的图像检索模型的训练方法的流程示意图;

[0016] 图2为本申请实施例一所提供的图像检索模型的训练方法的另一流程示意图;

[0017] 图3为本申请实施例二所提供的图像检索模型的训练方法的流程示意图;

[0018] 图4为本申请实施例二所提供的特征提取网络的结构图;

- [0019] 图5为本申请实施例二所提供的第一融合模块的结构图；
- [0020] 图6为本申请实施例三所提供的图像检索模型的训练方法的流程示意图；
- [0021] 图7为本申请实施例三所提供的特征提取网络的结构图；
- [0022] 图8为本申请实施例四所提供的图像检索方法的流程示意图；
- [0023] 图9为本申请实施例四所提供的图像检索方法的另一流程示意图；
- [0024] 图10为本申请实施例四所提供的ROI的处理方式的示例图；
- [0025] 图11为本申请实施例四所提供的图像检索结果的示例图；
- [0026] 图12为本申请实施例五所提供的图像检索模型的训练装置的结构示意图；
- [0027] 图13为本申请实施例六所提供的图像检索装置的结构示意图；
- [0028] 图14示出了适于用来实现本申请实施方式的示例性电子设备的框图。

### 具体实施方式

[0029] 下面详细描述本申请的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,旨在用于解释本申请,而不能理解为对本申请的限制。

[0030] 需要说明的是,本申请的技术方案中,所涉及的用户个人信息的获取,存储和应用等,均符合相关法律法规的规定,且不违背公序良俗。

[0031] 相关技术中的图像检索技术,在模型训练阶段一般是将图像中的ROI缩放至固定尺寸后直接输入网络,这种方式训练得到的模型,在进行检索时容易检索到与已知图像不相关的图像,检索结果的准确性差。

[0032] 本申请针对相关技术中训练得到的图像检索模型,在进行检索时容易检索到与已知图像不相关的图像,检索结果的准确性差的技术问题,提出一种图像检索模型的训练方法、图像检索方法、装置、电子设备、存储介质以及计算机程序产品。

[0033] 其中,图像检索模型的训练方法包括:获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注;将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象;将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。由此,能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0034] 下面参考附图描述本申请实施例的图像检索模型的训练方法、图像检索方法、装置、电子设备、存储介质以及计算机程序产品。

[0035] 首先对本申请实施例提供的图像检索模型的训练方法进行说明。

[0036] 其中,需要说明的是,本申请实施例提供的图像检索模型的训练方法,由图像检索模型的训练装置执行。其中,该图像检索模型的训练装置可以为电子设备,也可以被配置在电子设备中,以通过执行本申请实施例提供的图像检索模型的训练方法,实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0037] 其中,电子设备可以为个人电脑(Personal Computer,简称PC)、云端设备、移动设

备、服务器等,移动设备例如可以为手机、平板电脑、个人数字助理、穿戴式设备、车载设备等任意硬件设备,本申请对此不作限制。

[0038] 图1为本申请实施例一所提供的图像检索模型的训练方法的流程示意图。如图1所示,该图像检索模型的训练方法可以包括以下步骤101-104。

[0039] 步骤101,获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注。

[0040] 其中,样本图像,为用于训练图像检索模型的训练样本。样本图像中包括的样本对象,为样本图像中的主体对象,可以为商品、人物、动物、物品等。

[0041] 在示例实施例中,可以根据应用场景的需要,选择作为训练样本的样本图像。比如,在图像检索模型用于电商领域的同款推荐场景时,可以将包括商品的图像作为样本图像,用于对图像检索模型进行训练。

[0042] 需要说明的是,本申请实施例中训练图像检索模型时的上述训练数据,可以由用户提供并被授权使用,或者从公开数据集中获取,或者通过其它符合相关法律法规的规定的方式获取,本申请对此不作限制。

[0043] 另外,本申请实施例中,在对所有样本图像中的样本对象进行类别划分时,可以将同一样本对象划分为一个类别,即相同的样本对象属于相同的类别,不同的样本对象属于不同的类别。

[0044] 比如,以多个样本图像中包括的1000个样本图像为例,假设其中标号为1-200的样本图像是对物品A从多个角度进行拍摄得到的图像,其中标号为201-400的样本图像是对物品B从多个角度进行拍摄得到的图像,其中标号为401-650的样本图像是对物品C从多个角度进行拍摄得到的图像,其中651-1000的样本图像是对物品D从多个角度进行拍摄得到的图像。则该1000个样本图像中样本对象的所属类别包括4类,假设物品A的所属类别为类别a,物品B的所属类别为类别b,物品C的所属类别为类别c,物品D的所属类别为类别d。则标号为1-200的样本图像以类别a进行标注,标号为201-400的样本图像以类别b进行标注,标号为401-650的样本图像以类别c进行标注,标号为651-1000的样本图像以类别d进行标注。

[0045] 步骤102,将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象。

[0046] 其中,图像检索模型,为具有图像检索功能的神经网络模型。其中,图像检索模型中包括依次连接的目标检测网络和特征提取网络。

[0047] 其中,目标检测网络,用于检测任意图像中的对象,以获取该任意图像中的ROI和该ROI的尺寸信息,其中该ROI中包括该任意图像中的主体对象。ROI的尺寸信息,可以包括ROI的高度和宽度。

[0048] 本申请实施例中,参考图2,对于每个样本图像,可以将该样本图像输入图像检索模型中的目标检测网络21,目标检测网络21输出的ROI,即为该样本图像中包括样本对象的样本图像块,同时,目标检测网络21还可以输出样本图像块的尺寸信息。

[0049] 步骤103,将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量。

[0050] 其中,特征提取网络,用于基于任意图像中的ROI以及ROI的尺寸信息,提取该任意图像中的主体对象的特征向量,其中,该特征向量表征主体对象的属性特征,比如商品的颜

色、花纹、款式、大小等属性特征。

[0051] 本申请实施例中,参考图2,对于每个样本图像,可以将该样本图像中的样本图像块(ROI)以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络22,特征提取网络22即可输出该样本图像中样本对象的特征向量,本申请实施例中为了便于与其它特征向量进行区分,将该特征向量称为预测特征向量。

[0052] 步骤104,基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。

[0053] 在本申请的实施例中,参考图2,可以根据特征提取网络22输出的各样本图像中样本对象的预测特征向量进行损失计算。具体的,可以根据特征提取网络22输出的各样本图像中样本对象的预测特征向量,对各样本图像中样本对象的所属类别进行预测,并将预测得到的所属类别以及标注的所属类别,代入损失函数,以确定损失值,进而根据损失值调整图像检索模型中目标检测网络与特征提取网络的模型参数,通过多次迭代优化,得到训练后的图像检索模型。

[0054] 其中,损失函数,可以根据需要设置,比如可以设置为交叉熵损失函数、均方误差MSE(mean squared error)损失函数,或者其它损失函数,本申请对此不作限制。

[0055] 本申请实施例中,在对图像检索模型进行训练时,由于将样本图像以及样本图像中样本图像块的尺寸信息同时输入特征提取网络进行模型训练,相比将图像的ROI缩放至固定尺寸后直接输入网络,本申请实施例的方案可以使得图像检索模型学习到图像中主体对象的真实形态,比如衣服的长短款,进而使得在利用图像检索模型进行图像检索时,可以提高检索结果的准确性。

[0056] 综上,本申请实施例提供的图像检索模型的训练方法,获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注,将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象,将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量,基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。由此,能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0057] 在一种可能的实现形式中,特征提取网络可以包括第一尺寸特征提取模块、图像特征提取模块,以及与第一尺寸特征提取模块和图像特征提取模块连接的第一融合模块。下面针对上述结构的特征提取网络,结合图3,对本申请实施例中的图像检索模型的训练方法中,将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量的过程进行进一步说明。

[0058] 图3为本申请实施例二所提供的图像检索模型的训练方法的流程示意图。如图3所示,图像检索模型的训练方法可以包括以下步骤301-306。

[0059] 步骤301,获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注。

[0060] 步骤302,将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象。

[0061] 其中,步骤301-302的具体实现过程及原理,可以参考其它实施例的描述,此处不再赘述。

[0062] 步骤303,对于每个样本图像,将其中的样本图像块的尺寸信息,输入特征提取网络包括的第一尺寸特征提取模块,以基于尺寸信息获取样本图像块的第一尺寸特征。

[0063] 参考图4,特征提取网络可以包括第一尺寸特征提取模块41、图像特征提取模块42,以及与第一尺寸特征提取模块41和图像特征提取模块42连接的第一融合模块43。其中,第一尺寸特征提取模块41用于提取任意图像块的尺寸特征。图像特征提取模块42用于提取任意图像块的图像特征向量。第一融合模块43,用于将尺寸特征和图像特征向量进行融合,得到图像中主体对象的预测特征向量。

[0064] 本申请实施例中,对于每个样本图像,将其中的样本图像块的尺寸信息,输入第一尺寸特征提取模块41,可以得到样本图像块的第一尺寸特征。

[0065] 其中,样本图像块的尺寸信息,可以包括样本图像块的宽度和高度。其中,以 $I_{iw}$ 表示第 $i$ 个样本图像块的宽度,以 $I_{ih}$ 表示第 $i$ 个样本图像块的高度。

[0066] 在一种可能的实现形式中,对于第 $i$ 个样本图像块,可以直接将该样本图像块的宽度 $I_{iw}$ 的值和高度 $I_{ih}$ 的值,确定为该样本图像块的第一尺寸特征。在该情况下,第一尺寸特征的维度为2。

[0067] 在另一种可能的实现形式中,对于第 $i$ 个样本图像块,可以将该样本图像块的宽度和高度的比值 $r_i = \frac{I_{iw}}{I_{ih}}$ ,即孔径比,确定为该样本图像块的第一尺寸特征。在该情况下,第一尺寸特征的维度为1。

[0068] 在另一种可能的实现形式中,可以分别计算各样本图像中样本图像块的宽度和高度的比值,并确定各比值的均值 $\mu_r$ 和标准差 $\sigma_r$ ,进而对于第 $i$ 个样本图像块,可以通过以下公式(1),确定该样本图像块的第一尺寸特征。在该情况下,第一尺寸特征的维度为1。

$$[0069] \quad f_{ratio} = \frac{r_i - \mu_r}{\sigma_r} \quad (1)$$

[0070] 其中, $f_{ratio}$ 表示第 $i$ 个样本图像块的第一尺寸特征; $r_i$ 表示第 $i$ 个样本图像块的宽度和高度的比值; $\mu_r$ 表示所有样本图像块的宽度和高度的比值的均值; $\sigma_r$ 表示所有样本图像块的宽度和高度的比值的标准差。

[0071] 在另一种可能的实现形式中,可以分别计算各样本图像中样本图像块的宽度和高度的对数比值,并确定各对数比值的均值 $\mu_{la}$ 和标准差 $\sigma_{la}$ ,进而对于第 $i$ 个样本图像块,可以通过以下公式(2),确定该第 $i$ 个样本图像块的第一尺寸特征。其中,对数比值,即比值取对数。在该情况下,第一尺寸特征的维度为1。

$$[0072] \quad f_{ratio} = \frac{\log_a r_i - \mu_{la}}{\sigma_{la}} \quad (2)$$

[0073] 其中, $f_{ratio}$ 表示第 $i$ 个样本图像块的第一尺寸特征; $\log_a r_i$ 表示第 $i$ 个样本图像块的宽度和高度的对数比值; $\mu_{la}$ 表示所有样本图像块的宽度和高度的对数比值的均值; $\sigma_{la}$ 表示所有样本图像块的宽度和高度的对数比值的标准差。

[0074] 步骤304,将样本图像块输入特征提取网络包括的图像特征提取模块,以对样本图像块进行图像特征提取,得到样本图像块的图像特征向量。

[0075] 在本申请实施例中,可以将样本图像块输入图像特征提取模块,以对样本图像块进行尺寸缩放,得到预设尺寸的目标样本图像块,并对预设尺寸的目标样本图像块进行图像特征提取,得到样本图像块中样本对象的图像特征向量。

[0076] 在一种可能的实现形式中,参考图4,图像特征提取模块42可以包括依次连接的图像处理子模块421以及第一主干网络422。其中,图像处理子模块421用于对任意图像块进行尺寸缩放,得到预设尺寸的图像块,并获取该图像块对应的张量。第一主干网络422,用于对任意图像块进行特征提取,得到图像特征向量,其可以为任意能够实现特征提取的网络,本申请对此不作限制。

[0077] 参考图4,对于任意样本图像块,可以将该样本图像块输入图像处理子模块421,以对样本图像块进行尺寸缩放,得到预设尺寸的第一图像块,并获取第一图像块对应的第一张量,进而将第一张量输入第一主干网络422,以基于第一张量进行图像特征提取,得到样本图像块的图像特征向量。

[0078] 其中,预设尺寸,可以根据需要设置,本申请对此不作限制。通常,可以将预设尺寸设置为宽度和高度相等,比如预设尺寸为24\*24像素。

[0079] 其中,第一张量可以包括样本图像块中各像素点在多个通道(比如红色通道、蓝色通道和绿色通道)的像素值。

[0080] 需要说明的是,为了使得第一主干网络可以实现对第一张量进行图像特征提取,需要第一主干网络中第一个卷积核的输入通道数与第一张量的通道数相等。比如,在第一张量的尺寸为24\*24\*3时,即第一张量的通道数为3时,第一主干网络中第一个卷积核的尺寸为[64,3,7,7]。其中,64表示卷积核的输出通道数,3表示卷积核的输入通道数。

[0081] 步骤305,将第一尺寸特征和图像特征向量,输入特征提取网络中与第一尺寸特征提取模块和图像特征提取模块连接的第一融合模块,以将第一尺寸特征和图像特征向量融合,得到样本图像中样本对象的预测特征向量。

[0082] 在一种可能的实现形式中,参考图5,第一融合模块43包括依次连接的多层感知机子模块51以及融合子模块52。其中,多层感知机子模块51可以对尺寸特征进行维度扩展,得到预设维度的尺寸特征。融合子模块52,可以对预设维度的尺寸特征和图像特征向量进行融合,得到预测特征向量。

[0083] 参考图5,对于任意样本图像块,可以将该样本图像块的第一尺寸特征输入多层感知机子模块51,以对第一尺寸特征进行维度扩展,得到第二尺寸特征,进而将第二尺寸特征和图像特征向量输入融合子模块52,以将第二尺寸特征和图像特征向量融合,得到样本图像中样本对象的预测特征向量。

[0084] 其中,可以通过以下多种方式,将第二尺寸特征和图像特征向量融合,得到样本图像中样本对象的预测特征向量。

[0085] 比如,可以将第二尺寸特征和图像特征向量拼接或相加,得到样本图像中样本对象的预测特征向量。或者,可以将第二尺寸特征和图像特征向量基于自注意力机制进行融合,得到样本图像中样本对象的预测特征向量。

[0086] 比如,假设第二尺寸特征为128维的特征,图像特征向量为2048维的特征,则将第二尺寸特征和图像特征向量拼接,可以得到2176维的预测特征向量。

[0087] 在本申请的实施例中,可以通过下表1所示的方式,设置多层感知机子模块51在不

同层数 $N_{layer}$ 下的输出特征尺寸(即维度)。其中,多层感知机子模块51的输入特征尺寸 $infeatdim=1$ ,输出特征尺寸 $outfeatdim \in [16,1024]$ 。

[0088] 表1多层感知机子模块的层数设计

层数 $N_{layer}$	输入特征尺寸	输出特征尺寸
1	1	$outfeatdim$
2	第1层: 1 第2层: $outfeatdim / 2$	第1层: $outfeatdim / 2$ 第2层: $outfeatdim$
3	第1层: 1 第2层: $outfeatdim / 4$ 第3层: $outfeatdim / 2$	第1层: $outfeatdim / 4$ 第2层: $outfeatdim / 2$ 第3层: $outfeatdim$

[0090] 步骤306,基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。

[0091] 其中,步骤306的具体实现过程及原理,可以参考其它实施例的描述,此处不再赘述。

[0092] 由此,能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0093] 在一种可能的实现形式中,特征提取网络可以包括第二尺寸特征提取模块、图像处理模块、与第二尺寸特征提取模块及图像处理模块连接的第二融合模块,以及与第二融合模块连接的第二主干网络。下面针对上述结构的特征提取网络,结合图6,对本申请实施例中的图像检索模型的训练方法中,将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量的过程进行进一步说明。

[0094] 图6为本申请实施例三所提供的图像检索模型的训练方法的流程示意图。如图6所示,图像检索模型的训练方法可以包括以下步骤601-607。

[0095] 步骤601,获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注。

[0096] 步骤602,将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象。

[0097] 其中,步骤601-602的具体实现过程及原理,可以参考其它实施例的描述,此处不再赘述。

[0098] 步骤603,对于每个样本图像,将样本图像中样本图像块的尺寸信息输入特征提取网络包括的第二尺寸特征提取模块,以基于尺寸信息获取样本图像块的第三尺寸特征。

[0099] 参考图7,特征提取网络可以包括第二尺寸特征提取模块71、图像处理模块72、与第二尺寸特征提取模块71及图像处理模块72连接的第二融合模块73,以及与第二融合模块73连接的第二主干网络74。其中,第二尺寸特征提取模块71用于提取任意图像块的尺寸特征。图像处理模块72用于对任意图像块进行尺寸缩放,得到预设尺寸的图像块,并获取该图

像块对应的张量。第二融合模块73,用于将尺寸特征和张量进行融合,得到融合特征向量。

[0100] 本申请实施例中,对于每个样本图像,将其中的样本图像块的尺寸信息,输入第二尺寸特征提取模块71,可以得到样本图像块的第三尺寸特征。

[0101] 其中,样本图像块的尺寸信息,可以包括样本图像块的宽度和高度。基于尺寸信息获取样本图像块的第三尺寸特征的方式,可以参考其它实施例中基于尺寸信息获取样本图像块的第一尺寸特征的方式,二者方式类似,此处不再赘述。

[0102] 步骤604,将样本图像块输入特征提取网络包括的图像处理模块,以对样本图像块进行尺寸缩放,得到预设尺寸的第二图像块,并获取第二图像块对应的第二张量。

[0103] 其中,预设尺寸,可以根据需要设置,本申请对此不作限制。通常,可以将预设尺寸设置为宽度和高度相等,比如预设尺寸为24\*24像素。

[0104] 其中,第二张量可以包括样本图像块中各像素点在多个通道(比如红色通道、蓝色通道和绿色通道)的像素值。

[0105] 参考图7,将样本图像块输入特征提取网络包括的图像处理模块72,可以对样本图像块进行尺寸缩放,得到预设尺寸的第二图像块,并获取第二图像块对应的第二张量。

[0106] 步骤605,将第三尺寸特征和第二张量,输入特征提取网络中与第二尺寸特征提取模块及图像处理模块连接的第二融合模块,以将第三尺寸特征和第二张量融合,得到融合特征向量。

[0107] 参考图7,将第三尺寸特征和第二张量,输入特征提取网络中与第二尺寸特征提取模块71及图像处理模块72连接的第二融合模块73,可以实现将第三尺寸特征和第二张量融合,得到融合特征向量。

[0108] 在本申请的实施例中,可以将第三尺寸特征和第二张量沿通道维度拼接,得到融合特征向量。

[0109] 具体的,可以通过以下方式,实现将第三尺寸特征和第二张量沿通道维度拼接,得到融合特征向量,即步骤605可以通过以下方式实现:

[0110] 将第三尺寸特征和第二张量,输入第二融合模块,以对于样本图像块中的每个像素点,将第三尺寸特征与像素点在多个通道的像素值拼接,得到像素点在多个通道的特征向量;基于样本图像块中的各个像素点在多个通道的特征向量,生成融合特征向量。

[0111] 举例来说,假设第三尺寸特征的维度为1。第二张量的尺寸为24\*24\*3,即第二张量包括样本图像块中24\*24个像素点在红色通道、蓝色通道和绿色通道这3个通道的像素值。则对于样本图像块中的每个像素点,可以将第三尺寸特征与像素点在3个通道的像素值拼接,得到像素点在4个通道的特征向量,其中,对于每个像素点来说,该像素点的特征向量包括红色通道、蓝色通道和绿色通道这3个通道的像素值以及第三尺寸特征。从而可以基于样本图像块中的24\*24个像素点在4个通道的特征向量,得到24\*24\*4维的融合特征向量。

[0112] 需要说明的是,由于图像的像素点的像素值的取值范围为 $[-128, 128]$ ,而在第三尺寸特征为一维时,第三尺寸特征的值为小数,若将小数形式的第三尺寸特征和像素点在多个通道的像素值拼接,则得到的融合特征向量可能会引起模型的波动,那么,可以将第三尺寸特征归一化到与像素值的取值范围接近,再与像素点在多个通道的像素值拼接,以保证模型的检索准确性。

[0113] 步骤606,将融合特征向量输入特征提取网络中与第二融合模块连接的第二主干

网络,以基于融合特征向量进行特征提取,得到样本图像中样本对象的预测特征向量。

[0114] 第二主干网络,用于对任意特性向量进行特征提取,得到预测特征向量,其可以为任意能够实现特征提取的网络,本申请对此不作限制。

[0115] 参考图7,将融合特征向量输入特征提取网络中与第二融合模块73连接的第二主干网络74,可以实现基于融合特征向量进行特征提取,得到样本图像中样本对象的预测特征向量。

[0116] 需要说明的是,为了使得第二主干网络可以实现对融合特征向量进行特征提取,需要第二主干网络中第一个卷积核的输入通道数与融合特征向量的通道数相等。比如,在融合特征向量为 $24*24*4$ 维时,即融合特征向量的通道数为4时,第二主干网络中第一个卷积核的尺寸为 $[64,4,7,7]$ 。其中,64表示卷积核的输出通道数,4表示卷积核的输入通道数。

[0117] 其中,第二主干网络中第一个卷积核可以采用正交的方式,实现权重的初始化,或者,也可以采用其它的方式,实现权重的初始化,本申请对此不作限制。

[0118] 步骤607,基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。

[0119] 其中,步骤607的具体实现过程及原理,可以参考其它实施例的描述,此处不再赘述。

[0120] 由此,能够实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0121] 基于上述实施例中的图像检索模型的训练方法,本申请实施例中提供一种图像检索方法。下面对本申请实施例提供的图像检索方法进行说明。

[0122] 其中,需要说明的是,本申请实施例提供的图像检索方法,由图像检索装置执行。其中,该图像检索装置可以为电子设备,也可以被配置在电子设备中,以通过执行本申请实施例提供的图像检索方法,提高图像检索的准确性。

[0123] 其中,电子设备可以为个人电脑(Personal Computer,简称PC)、云端设备、移动设备、服务器等,移动设备例如可以为手机、平板电脑、个人数字助理、穿戴式设备、车载设备等任意硬件设备,本申请对此不作限制。

[0124] 图8为本申请实施例四所提供的图像检索方法的流程示意图。如图8所示,该图像检索方法可以包括以下步骤801-804。

[0125] 步骤801,获取待检索的检索图像。

[0126] 其中,检索图像中包括目标对象,目标对象为检索图像中的主体对象,可以为商品、人物、动物、物品等。

[0127] 步骤802,将检索图像输入图像检索模型中的目标检测网络,以获取检索图像中的目标图像块以及目标图像块的尺寸信息,目标图像块包括检索图像中的目标对象。

[0128] 其中,图像检索模型,为具有图像检索功能的神经网络模型。其中,图像检索模型中包括依次连接的目标检测网络和特征提取网络。其中,图像检索模型基于前述任一实施例的图像检索模型的训练方法训练得到。

[0129] 其中,目标检测网络,用于检测任意图像中的对象,以获取该任意图像中的ROI和该ROI的尺寸信息,其中该ROI中包括该任意图像中的主体对象。

[0130] 在本申请实施例中,参考图9,将检索图像输入图像检索模型中的目标检测网络

21,目标检测网络21输出的ROI,即为该检索图像中包括目标对象的目标图像块,同时,目标检测网络21还可以输出目标图像块的尺寸信息。

[0131] 其中,目标图像块的尺寸信息,可以包括目标图像块的高度和宽度。

[0132] 步骤803,将检索图像中的目标图像块以及目标图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取目标对象的预测特征向量。

[0133] 其中,特征提取网络,用于基于任意图像中的ROI以及ROI的尺寸信息,提取该任意图像中的主体对象的特征向量,其中,该特征向量表征主体对象的属性特征,比如商品的颜色、花纹、款式、大小等属性特征。

[0134] 本申请实施例中,参考图9,将检索图像中的目标图像块以及目标图像块的尺寸信息,输入图像检索模型中的特征提取网络22,特征提取网络22即可输出该检索图像中目标对象的特征向量,本申请实施例中为了便于与其它特征向量进行区分,将该特征向量称为预测特征向量。

[0135] 其中,步骤802-803的具体实现过程,与模型训练过程中获取样本图像中的样本图像块、样本图像块的尺寸信息以及样本对象的预测特征向量的过程类似,此处不再赘述。

[0136] 步骤804,基于目标对象的预测特征向量,从多个候选图像中确定目标图像。

[0137] 在本申请的一个实施例中,步骤804可以通过以下方式得到:获取各候选图像中候选对象的候选特征向量;确定目标对象的预测特征向量与各候选图像中候选对象的候选特征向量之间的相似度;基于目标对象的预测特征向量与各候选图像中候选对象的候选特征向量之间的相似度,从多个候选图像中确定目标图像。

[0138] 其中,可以预设目标图像的数量,从而可以将与目标对象的预测特征向量的相似度最高的预设数量的候选图像,作为目标图像。

[0139] 参考图9,可以通过检索模块,计算目标对象的预测特征向量与各候选图像中候选对象的候选特征向量之间的相似度,以从多个候选图像中,确定目标图像。

[0140] 在本申请的一个实施例中,获取各候选图像中候选对象的候选特征向量之前,还包括:对于每个候选图像,将候选图像输入目标检测网络,以获取候选图像中的候选图像块以及候选图像块的尺寸信息,候选图像块包括候选图像中的候选对象;将候选图像中的候选图像块以及候选图像块的尺寸信息,输入特征提取网络,以获取候选对象的预测特征向量。

[0141] 其中,获取候选图像中的候选图像块、候选图像块的尺寸信息以及候选对象的预测特征向量的过程,与模型训练过程中获取样本图像中的样本图像块、样本图像块的尺寸信息以及样本对象的预测特征向量的过程类似,此处不再赘述。

[0142] 可以理解的是,相关技术中的图像检索技术,在模型训练阶段一般是将图像中的ROI缩放至固定尺寸的正方形后直接输入网络(以下将该方法称为方法1),这种方式训练得到的模型,在进行检索时容易检索到与已知图像(即检索图像)不相关的图像,检索结果的准确性差。其中,方法1对ROI的处理方式可以参考图10中的第三列图像。

[0143] 另外,相关技术为了提高图像检索模型的检索准确性,还提出以下方法2、方法3和方法3。其中,方法2是在数据处理阶段保持ROI的孔径比,其余部分使用0像素进行填充,从而得到正方形的图像,再根据正方形的图像进行特征提取,由此,使得输入到网络的图像可以体现其孔径比信息,以提高图像检索模型的检索准确性;方法3是在数据处理阶段保持

ROI的孔径比,其余部分使用检索前的原始图像进行填充,从而得到正方形的图像,再根据正方形的图像进行特征提取,由此,使得输入到网络的图像可以体现其孔径比信息,以提高图像检索模型的检索准确性;方法4是在模型训练阶段将图像中的ROI缩放至固定尺寸的正方形后直接输入网络,在图像检索阶段,将孔径比特征与检索图像的图像特征融合,使用融合后的特征进行检索,以提高图像检索模型的检索准确性。其中,方法2对ROI的处理方式可以参考图10中的第四列图像;方法3对ROI的处理方式可以参考图10中的第五列图像。

[0144] 方法2和方法3的方式,由于用于进行特征提取的正方形图像,使用0像素或原始图像进行填充,降低了图像分辨率,丢失了更多细节信息,从而会导致模型提取出的图像特征不准确,模型的检索准确性较差。方法4的方式,由于模型训练阶段和图像检索阶段,图像检索模型的特征尺度不同,从而在将训练好的图像检索模型用于图像检索时,需要对特征尺度进行调节,而调试过程相对繁琐,增加了上线工作量,且模型的检索准确性也较差。

[0145] 而本申请实施例提供的方法,模型训练阶段和图像检索阶段,图像检索模型的特征尺度相同,模型训练阶段训练好的图像检索模型,可以直接用于图像检索阶段,从而在提高模型的检索准确性的同时,避免了对特征尺度进行调节,减少了上线工作量。另外,本申请图3和图6所示实施例提供的图像检索模型的训练方法,与方法2和方法3相比,未保持ROI的孔径比,而是将ROI缩放至预设尺寸,从而能够最大程度的保留图像的细节信息,提高图像检索模型的检索准确性。

[0146] 参考下表2和图11,其中表2为在31万数据集上进行图像检索时,本申请实施例提供的图像检索方法与相关技术中的图像检索方法的检索结果的对比数据,图11为根据本申请实施例提供的图像检索方法进行图像检索得到的目标图像与利用方法1训练得到的模型进行图像检索得到的目标图像的对比示意图。

[0147] 表2实验结果

方法	准确率
方法1	95.94
方法2	95.56
方法3	95.14
图3所示方法	96.24
图6所示方法	96.29

[0149] 根据表2和图11可知,本申请实施例提供的图像检索方法与相关技术相比,能够提高检索准确性。

[0150] 综上,本申请实施例提供的图像检索方法,获取待检索的检索图像,将检索图像输入图像检索模型中的目标检测网络,以获取检索图像中的目标图像块以及目标图像块的尺寸信息,目标图像块包括检索图像中的目标对象,将检索图像中的目标图像块以及目标图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取目标对象的预测特征向量,基于目标对象的预测特征向量,从多个候选图像中确定目标图像,由此,提高了图像检索的准确性。

[0151] 图12为本申请实施例五所提供的图像检索模型的训练装置的结构示意图。

[0152] 如图12所示,该图像检索模型的训练装置1200可以包括:第一获取模块1210、第一处理模块1220、第二处理模块1230以及模型参数调整模块1240。

[0153] 其中,第一获取模块1210,用于获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注;

[0154] 第一处理模块1220,用于将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象;

[0155] 第二处理模块1230,用于将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量;

[0156] 模型参数调整模块1240,用于基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。

[0157] 需要说明的是,本申请实施例提供的图像检索模型的训练装置,可以执行前述实施例的图像检索模型的训练方法,该图像检索模型的训练装置可以为电子设备,也可以被配置在电子设备中,以通过执行本申请实施例提供的图像检索模型的训练方法,实现对图像检索模型的训练,且该图像检索模型在用于图像检索时的检索准确性高。

[0158] 其中,电子设备可以为PC、云端设备、移动设备、服务器等,移动设备例如可以为手机、平板电脑、个人数字助理、穿戴式设备、车载设备等任意硬件设备,本申请对此不作限制。

[0159] 在本申请实施例的一种可能的实现方式中,第二处理模块1230,包括:

[0160] 第一尺寸特征提取单元,用于对于每个样本图像,将其中的样本图像块的尺寸信息,输入特征提取网络包括的第一尺寸特征提取模块,以基于尺寸信息获取样本图像块的第一尺寸特征;

[0161] 图像特征提取单元,用于将样本图像块输入特征提取网络包括的图像特征提取模块,以对样本图像块进行图像特征提取,得到样本图像块的图像特征向量;

[0162] 第一融合单元,用于将第一尺寸特征和图像特征向量,输入特征提取网络中与第一尺寸特征提取模块和图像特征提取模块连接的第一融合模块,以将第一尺寸特征和图像特征向量融合,得到样本图像中样本对象的预测特征向量。

[0163] 在本申请实施例的一种可能的实现方式中,图像特征提取模块包括依次连接的图像处理子模块以及第一主干网络;

[0164] 图像特征提取单元,用于:

[0165] 将样本图像块输入图像处理子模块,以对样本图像块进行尺寸缩放,得到预设尺寸的第一图像块,并获取第一图像块对应的第一张量;

[0166] 将第一张量输入第一主干网络,以基于第一张量进行图像特征提取,得到样本图像块的图像特征向量。

[0167] 在本申请实施例的一种可能的实现方式中,第一融合模块包括依次连接的多层感知机子模块以及融合子模块;

[0168] 第一融合单元,用于:

[0169] 将第一尺寸特征输入多层感知机子模块,以对第一尺寸特征进行维度扩展,得到第二尺寸特征;

[0170] 将第二尺寸特征和图像特征向量输入融合子模块,以将第二尺寸特征和图像特征向量融合,得到样本图像中样本对象的预测特征向量。

[0171] 在本申请实施例的一种可能的实现方式中,尺寸信息包括样本图像块的宽度和高度;第一尺寸特征提取单元,用于:

[0172] 将样本图像块的宽度和高度,确定为第一尺寸特征;

[0173] 或者,将样本图像块的宽度和高度的比值,确定为第一尺寸特征;

[0174] 或者,确定各样本图像中样本图像块的宽度和高度的比值,并确定各比值的均值和标准差,基于比值、均值和标准差,确定第一尺寸特征;

[0175] 或者,确定各样本图像中样本图像块的宽度和高度的对数比值,并确定各对数比值的均值和标准差,基于对数比值、各对数比值的均值和标准差,确定第一尺寸特征。

[0176] 在本申请实施例的一种可能的实现方式中,第二处理模块1230,包括:

[0177] 第二尺寸特征提取单元,用于对于每个样本图像,将样本图像中样本图像块的尺寸信息输入特征提取网络包括的第二尺寸特征提取模块,以基于尺寸信息获取样本图像块的第三尺寸特征;

[0178] 图像处理单元,用于将样本图像块输入特征提取网络包括的图像处理模块,以对样本图像块进行尺寸缩放,得到预设尺寸的第二图像块,并获取第二图像块对应的第二张量;

[0179] 第二融合单元,用于将第三尺寸特征和第二张量,输入特征提取网络中与第二尺寸特征提取模块及图像处理模块连接的第二融合模块,以将第三尺寸特征和第二张量融合,得到融合特征向量;

[0180] 融合特征提取单元,用于将融合特征向量输入特征提取网络中与第二融合模块连接的第二主干网络,以基于融合特征向量进行特征提取,得到样本图像中样本对象的预测特征向量。

[0181] 在本申请实施例的一种可能的实现方式中,第二张量包括样本图像块中各像素点在多个通道的像素值;

[0182] 第二融合单元,用于:

[0183] 将第三尺寸特征和第二张量,输入第二融合模块,以对于样本图像块中的每个像素点,将第三尺寸特征与像素点在多个通道的像素值拼接,得到像素点在多个通道的特征向量;

[0184] 基于样本图像块中的各个像素点在多个通道的特征向量,生成融合特征向量。

[0185] 需要说明的是,前述图像检索模型的训练方法实施例中的解释说明也适用于该实施例的图像检索模型的训练装置,此处不再赘述。

[0186] 本申请实施例的图像检索模型的训练装置,获取多个样本图像,各样本图像以其中包括的样本对象的所属类别进行标注,将各样本图像输入图像检索模型中的目标检测网络,以获取对应样本图像中的样本图像块以及样本图像块的尺寸信息,样本图像块包括样本图像中的样本对象,将各样本图像中的样本图像块以及样本图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取对应样本图像中样本对象的预测特征向量,基于各样本图像中样本对象的预测特征向量以及样本对象的所属类别,确定损失值,并基于损失值,调整目标检测网络与特征提取网络的模型参数。由此,能够实现对图像检索模型的训

练,且该图像检索模型在用于图像检索时的检索准确性高。

[0187] 图13为本申请实施例六所提供的图像检索装置的结构示意图。

[0188] 如图13所示,该图像检索装置1300可以包括:第二获取模块1310、第三处理模块1320、第四处理模块1330以及确定模块1340。

[0189] 其中,第二获取模块1310,用于获取待检索的检索图像;

[0190] 第三处理模块1320,用于将检索图像输入图像检索模型中的目标检测网络,以获取检索图像中的目标图像块以及目标图像块的尺寸信息,目标图像块包括检索图像中的目标对象;

[0191] 第四处理模块1330,用于将检索图像中的目标图像块以及目标图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取目标对象的预测特征向量,其中,图像检索模型基于前述任一图像检索模型的训练方法训练得到;

[0192] 确定模块1340,用于基于目标对象的预测特征向量,从多个候选图像中确定目标图像。

[0193] 在本申请实施例的一种可能的实现方式中,确定模块1340,包括:

[0194] 获取单元,用于获取各候选图像中候选对象的候选特征向量;

[0195] 第一确定单元,用于确定目标对象的预测特征向量与各候选图像中候选对象的候选特征向量之间的相似度;

[0196] 第二确定单元,用于基于目标对象的预测特征向量与各候选图像中候选对象的候选特征向量之间的相似度,从多个候选图像中确定目标图像。

[0197] 在本申请实施例的一种可能的实现方式中,图像检索装置1300还可以包括:

[0198] 第三获取模块,用于对于每个候选图像,将候选图像输入目标检测网络,以获取候选图像中的候选图像块以及候选图像块的尺寸信息,候选图像块包括候选图像中的候选对象;

[0199] 第五处理模块,用于将候选图像中的候选图像块以及候选图像块的尺寸信息,输入特征提取网络,以获取候选对象的预测特征向量。

[0200] 需要说明的是,本申请实施例提供的图像检索装置,可以执行前述实施例的图像检索方法,该图像检索装置可以为电子设备,也可以被配置在电子设备中,以通过执行本申请实施例提供的图像检索方法,提高图像检索的准确性。

[0201] 其中,电子设备可以为PC、云端设备、移动设备、服务器等,移动设备例如可以为手机、平板电脑、个人数字助理、穿戴式设备、车载设备等任意硬件设备,本申请对此不作限制。

[0202] 需要说明的是,前述图像检索方法实施例中的解释说明也适用于该实施例的图像检索装置,此处不再赘述。

[0203] 本申请实施例的图像检索装置,获取待检索的检索图像,将检索图像输入图像检索模型中的目标检测网络,以获取检索图像中的目标图像块以及目标图像块的尺寸信息,目标图像块包括检索图像中的目标对象,将检索图像中的目标图像块以及目标图像块的尺寸信息,输入图像检索模型中的特征提取网络,以获取目标对象的预测特征向量,基于目标对象的预测特征向量,从多个候选图像中确定目标图像,由此,提高了图像检索的准确性。

[0204] 为了实现上述实施例,本申请还提出一种电子设备,包括:至少一个处理器;以及

与前述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如本申请前述任一实施例提出的图像检索模型的训练方法,或者执行如本申请前述任一实施例提出的图像检索方法。

[0205] 为了实现上述实施例,本申请还提出一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行如本申请前述任一实施例提出的图像检索模型的训练方法,或者执行如本申请前述任一实施例提出的图像检索方法。

[0206] 为了实现上述实施例,本申请还提出一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现如本申请前述任一实施例提出的图像检索模型的训练方法,或者实现如本申请前述任一实施例提出的图像检索方法。

[0207] 图14示出了适于用来实现本申请实施方式的示例性电子设备的框图。图14显示的电子设备1400仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0208] 如图14所示,电子设备1400以通用计算设备的形式表现。电子设备1400的组件可以包括但不限于:一个或者多个处理器或者处理单元16,系统存储器28,连接不同系统组件(包括系统存储器28和处理单元16)的总线18。

[0209] 总线18表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(Industry Standard Architecture;以下简称:ISA)总线,微通道体系结构(Micro Channel Architecture;以下简称:MAC)总线,增强型ISA总线、视频电子标准协会(Video Electronics Standards Association;以下简称:VESA)局域总线以及外围组件互连(Peripheral Component Interconnection;以下简称:PCI)总线。

[0210] 电子设备1400典型地包括多种计算机系统可读介质。这些介质可以是任何能够被电子设备1400访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0211] 存储器28可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(Random Access Memory;以下简称:RAM) 30和/或高速缓存存储器32。电子设备1400可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统34可以用于读写不可移动的、非易失性磁介质(图14未显示,通常称为“硬盘驱动器”)。尽管图14中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如:光盘只读存储器(Compact Disc Read Only Memory;以下简称:CD-ROM)、数字多功能只读光盘(Digital Video Disc Read Only Memory;以下简称:DVD-ROM) 或者其他光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线18相连。存储器28可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本申请各实施例的功能。

[0212] 具有一组(至少一个)程序模块42的程序/实用工具40,可以存储在例如存储器28中,这样的程序模块42包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块42通常执行本申请所描述的实施例中的功能和/或方法。

[0213] 电子设备1400也可以与一个或多个外部设备14(例如键盘、指向设备、显示器24等)通信,还可与一个或者多个使得用户能与该电子设备1400交互的设备通信,和/或与使得该电子设备1400能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口22进行。并且,电子设备1400还可以通过网络适配器20与一个或者多个网络(例如局域网(Local Area Network;以下简称:LAN),广域网(Wide Area Network;以下简称:WAN)和/或公共网络,例如因特网)通信。如图所示,网络适配器20通过总线18与电子设备1400的其它模块通信。应当明白,尽管图14中未示出,可以结合电子设备1400使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0214] 处理单元16通过运行存储在系统存储器28中的程序,从而执行各种功能应用以及数据处理,例如实现前述实施例中提及的方法。

[0215] 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。

[0216] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本申请的描述中,“多个”的含义是至少两个,例如两个,三个等,除非另有明确具体的限定。

[0217] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本申请的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本申请的实施例所属技术领域的技术人员所理解。

[0218] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为用于实现逻辑功能的可执行指令的定序列表,可以具体实现在任何计算机可读介质中,以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用,或结合这些指令执行系统、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPROM或闪速存储器),光纤装置,以及便携式光盘只读存储器(CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0219] 应当理解,本申请的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如,如果用硬件来实现和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列(PGA),现场可编程门阵列(FPGA)等。

[0220] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0221] 此外,在本申请各个实施例中的各功能单元可以集成在一个处理模块中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个模块中。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。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。

[0222] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。尽管上面已经示出和描述了本申请的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本申请的限制,本领域的普通技术人员在本申请的范围内可以对上述实施例进行变化、修改、替换和变型。

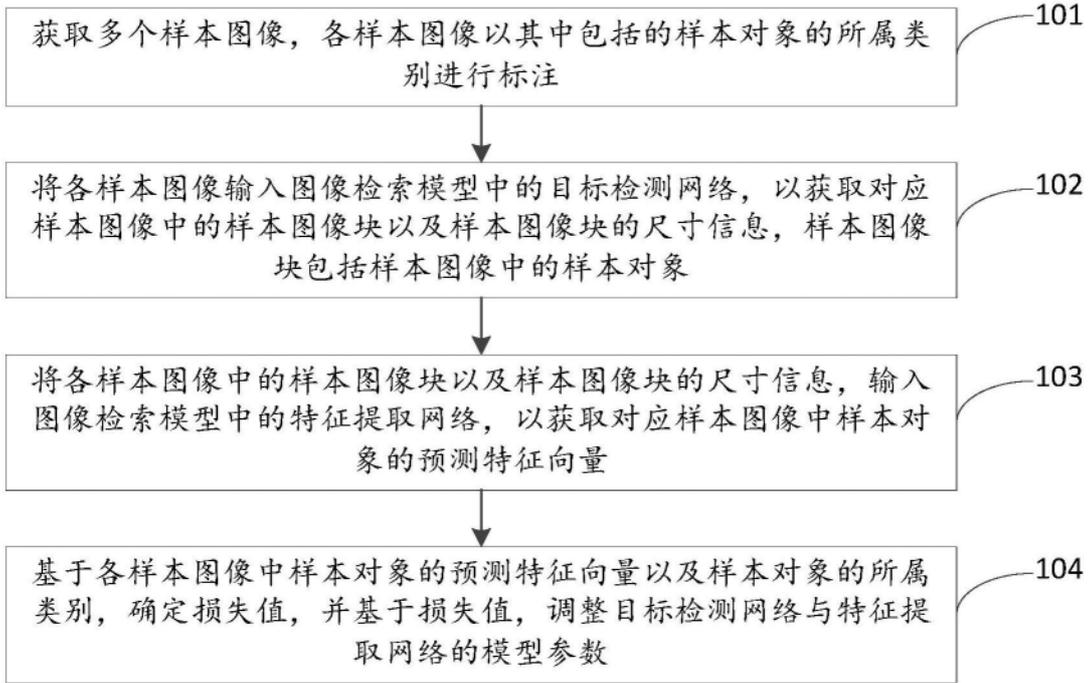


图1

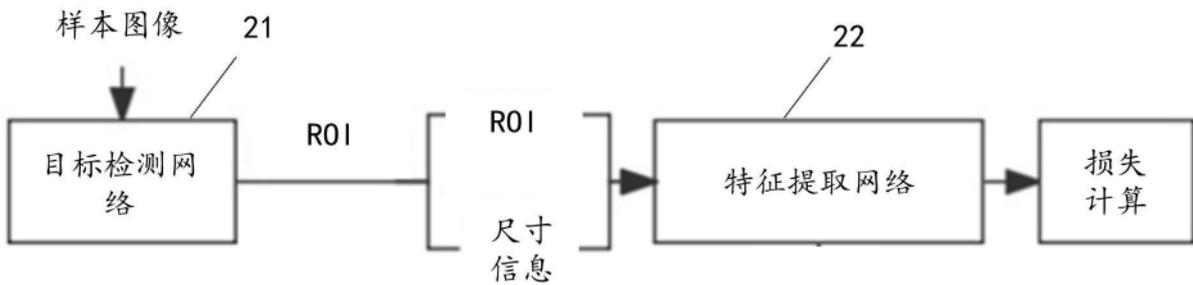


图2

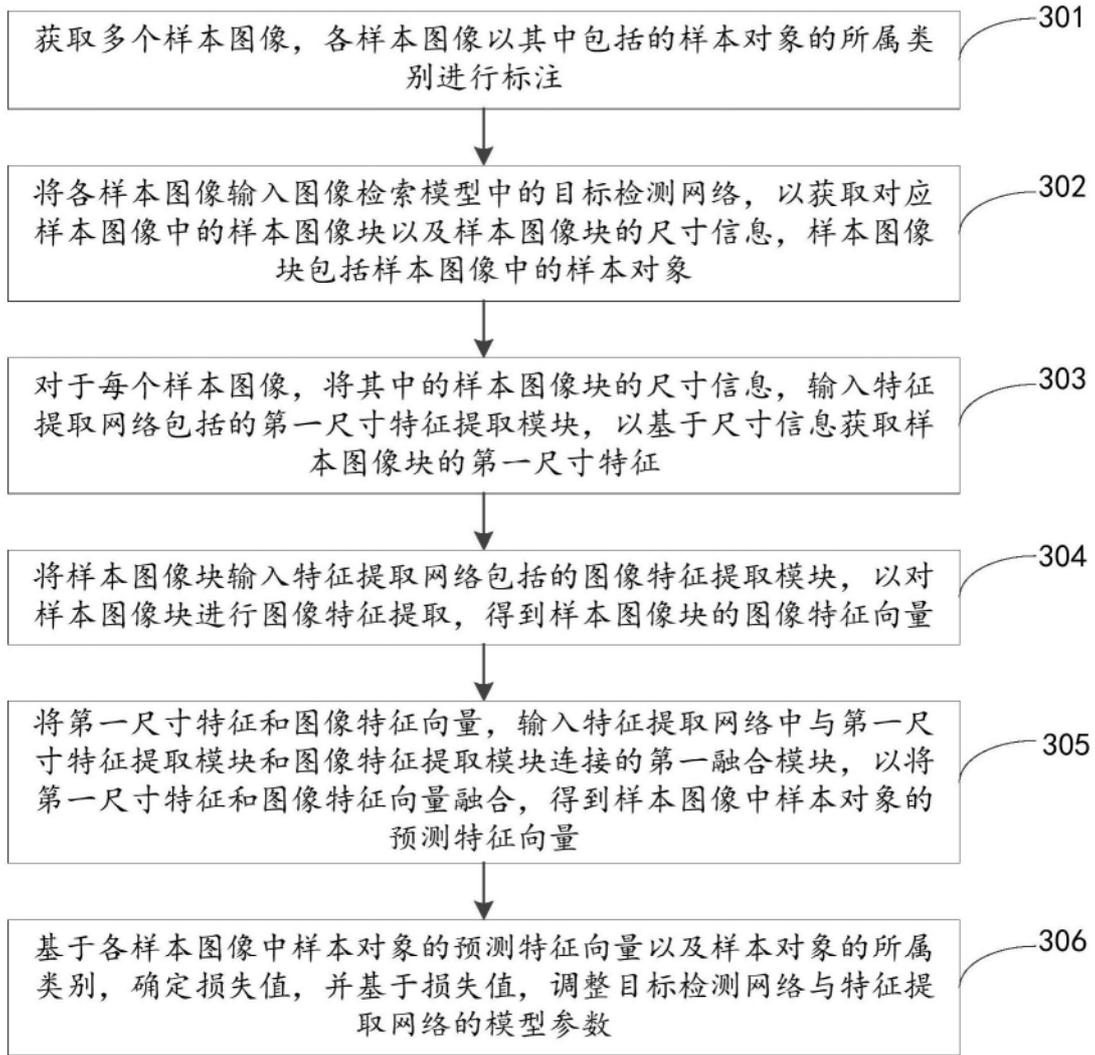


图3

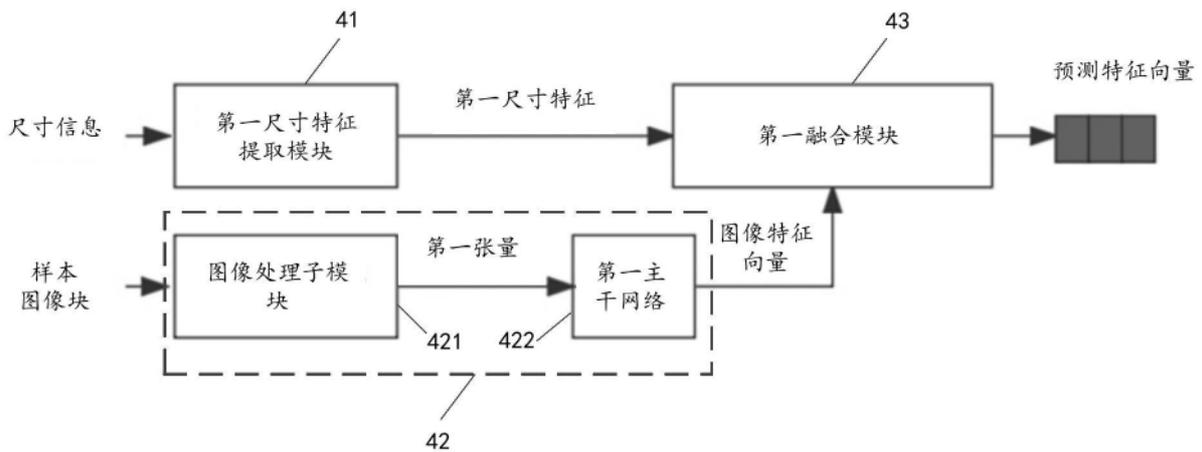


图4

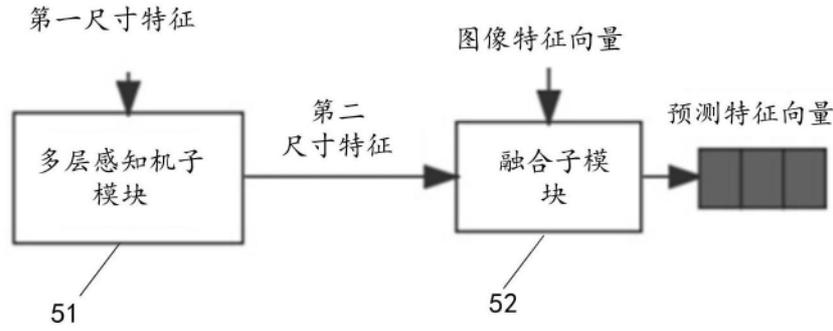


图5

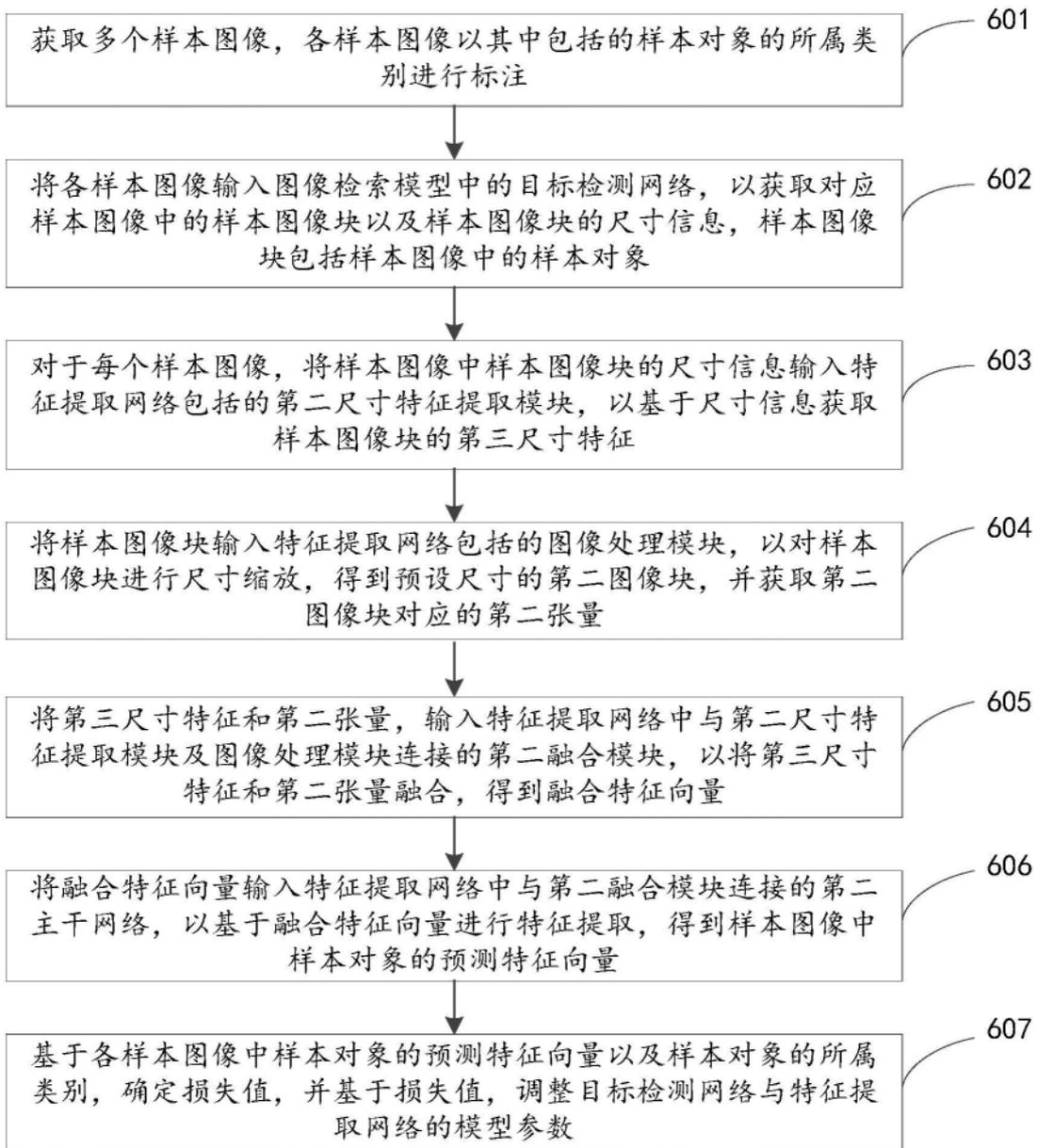


图6

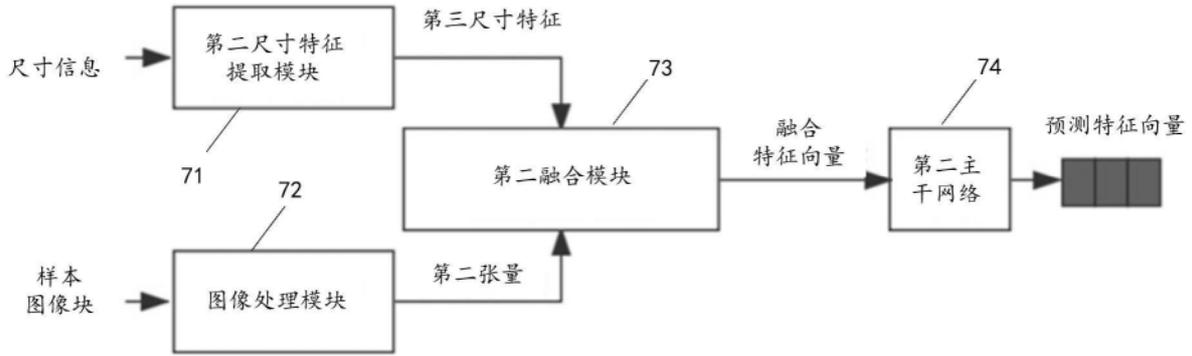


图7

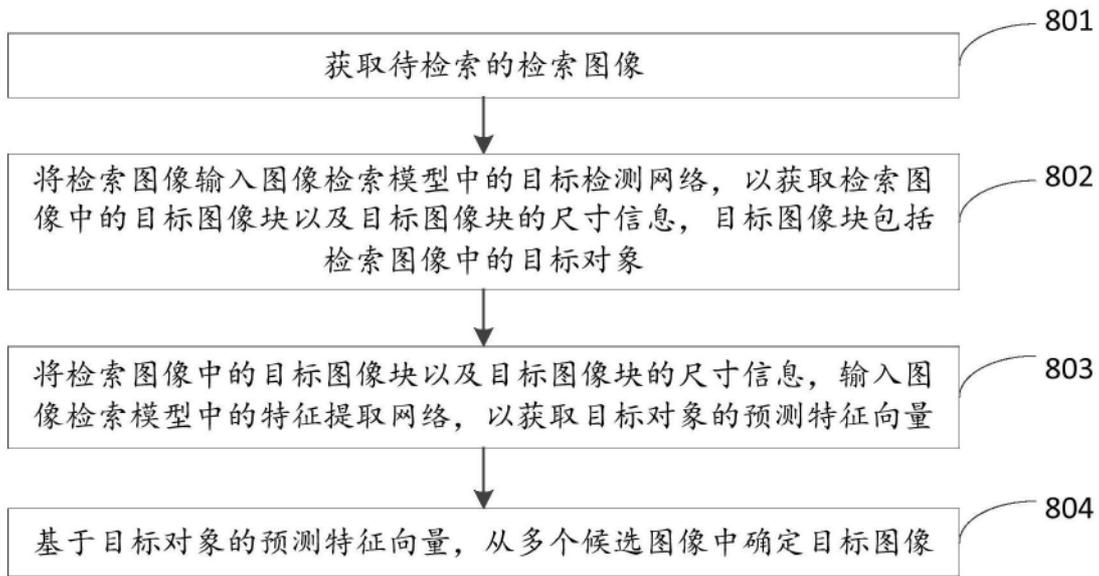


图8

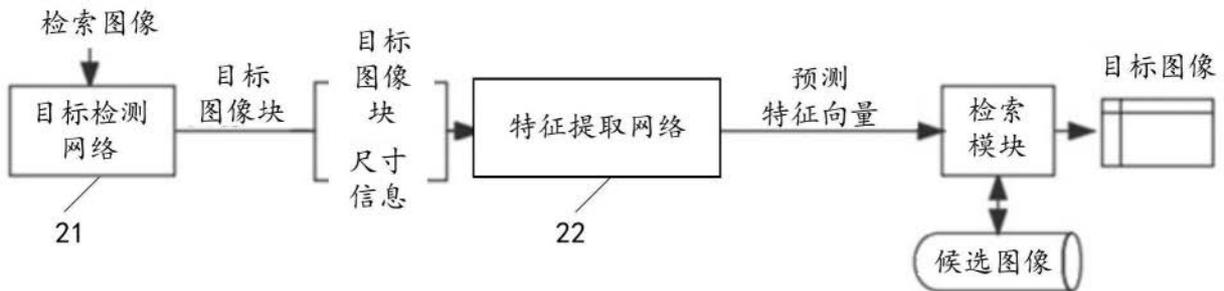


图9

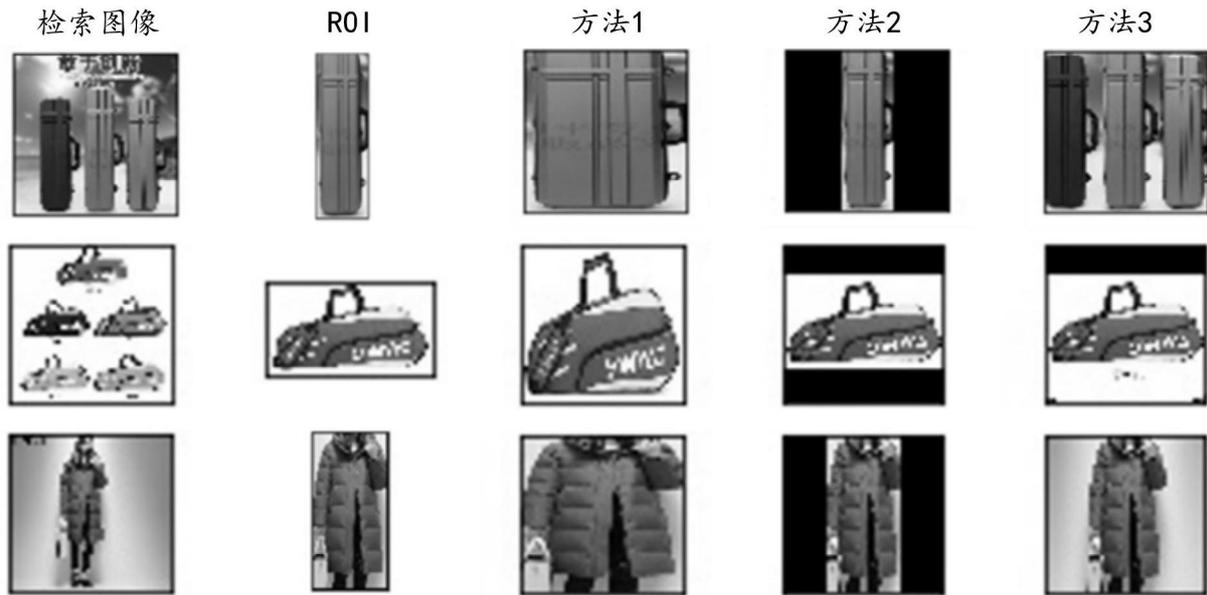


图10

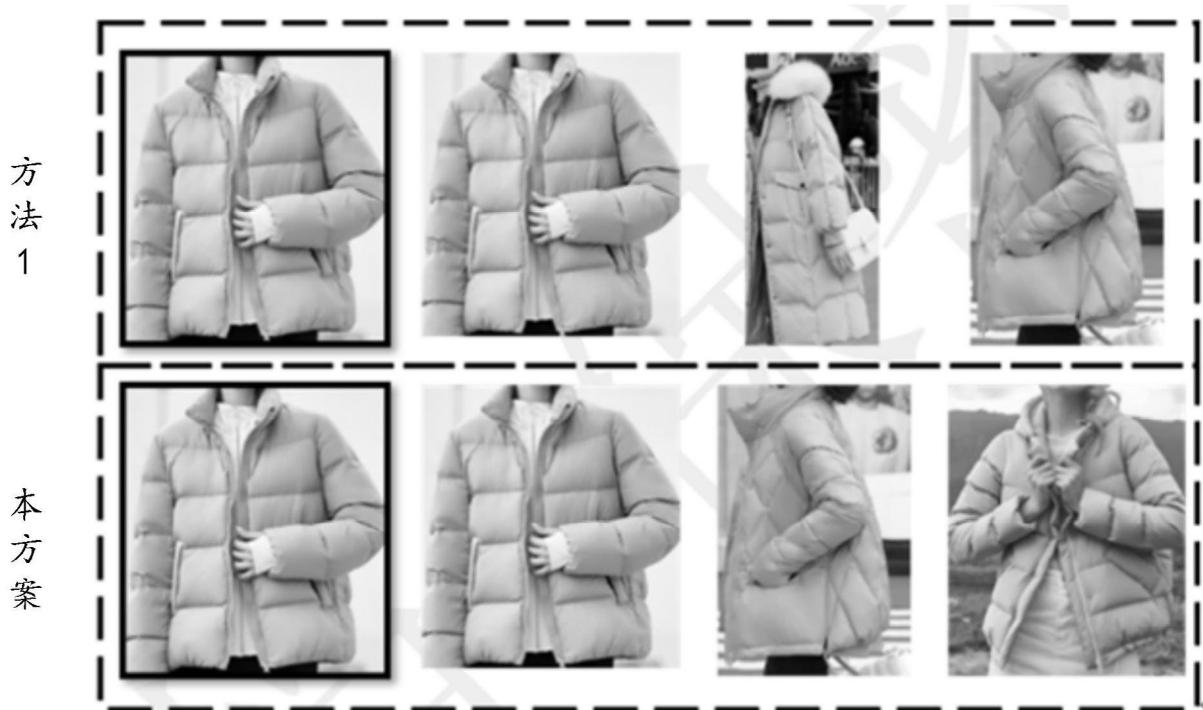


图11

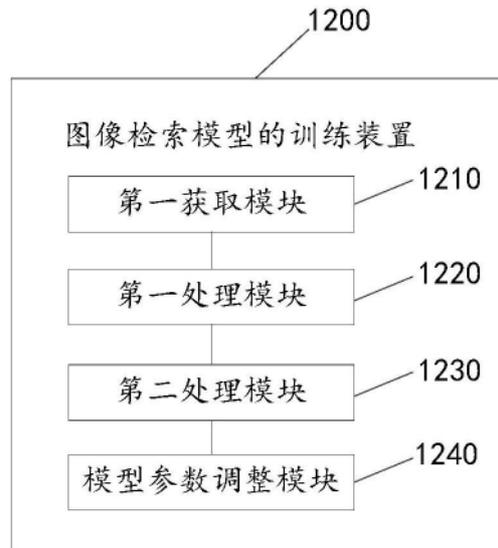


图12

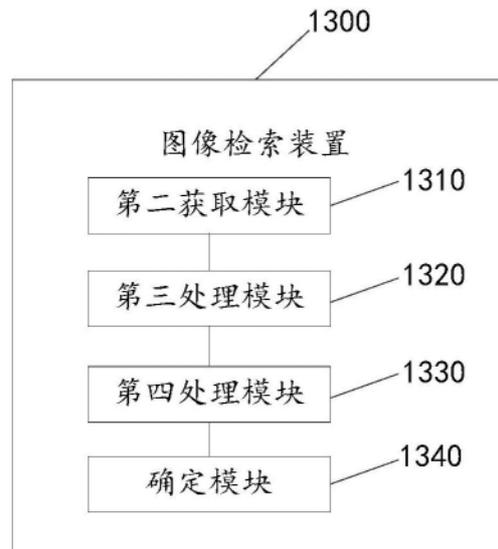


图13

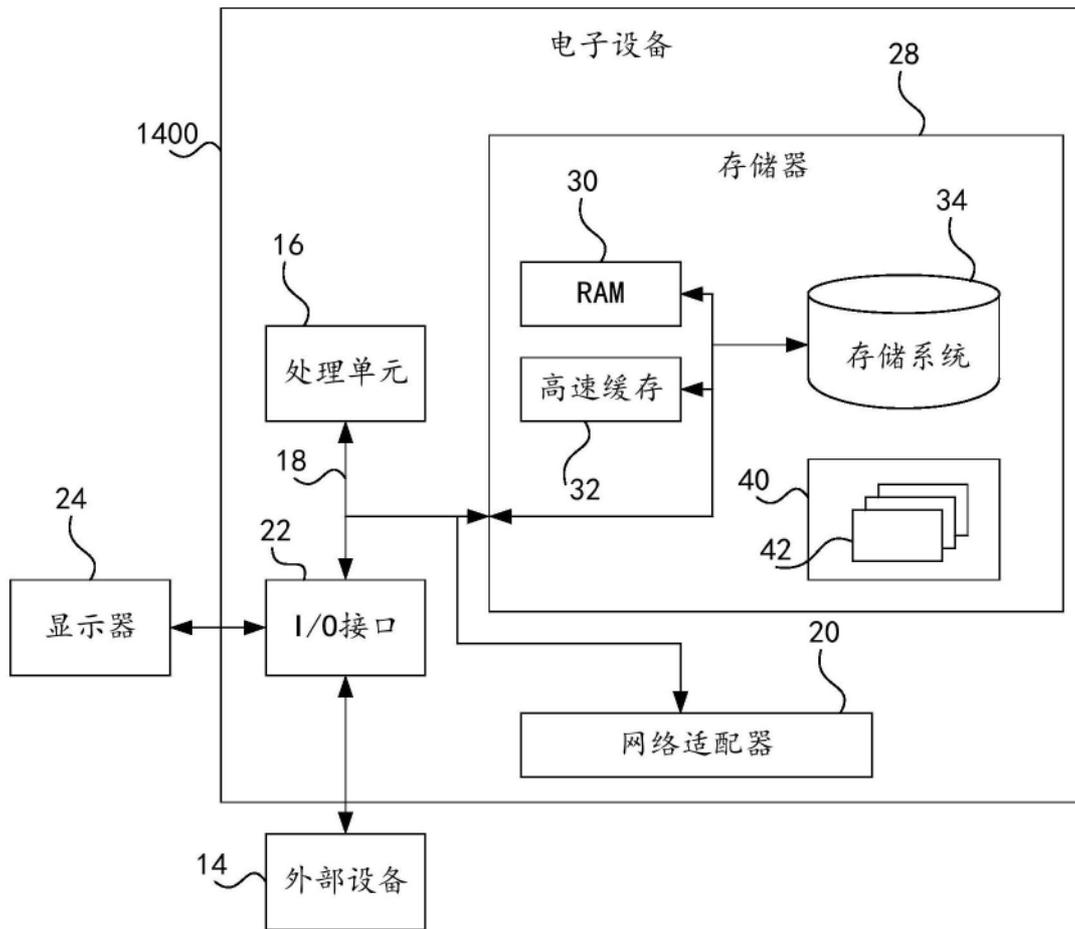


图14