



(12) 发明专利申请

(10) 申请公布号 CN 113505256 A

(43) 申请公布日 2021. 10. 15

(21) 申请号 202110747865.4

(22) 申请日 2021.07.02

(71) 申请人 北京达佳互联信息技术有限公司
地址 100085 北京市海淀区上地西路6号1
幢1层101D1-7

申请人 北京中科研究院

(72) 发明人 张磊 李攀登 李岩 袁勇
谢洪涛 张勇东

(74) 专利代理机构 北京三高永信知识产权代理
有限责任公司 11138

代理人 宁立存

(51) Int. Cl.

G06F 16/55 (2019.01)

G06F 16/583 (2019.01)

G06K 9/62 (2006.01)

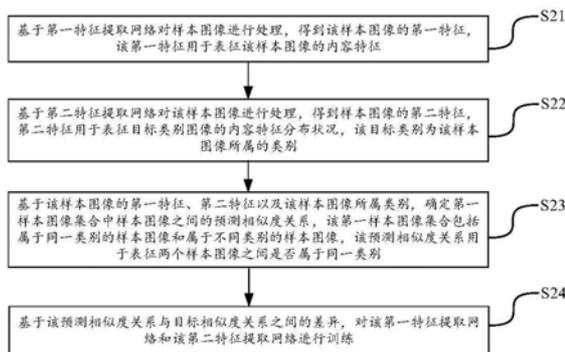
权利要求书3页 说明书36页 附图7页

(54) 发明名称

特征提取网络训练方法、图像处理方法及装置

(57) 摘要

本公开关于一种特征提取网络训练方法、图像处理方法及装置,涉及图像处理领域。本公开实施例,新增一种用于提取第二特征的第二特征提取网络,第二特征用于表征目标类别图像的内容特征分布状况,也即是该样本图像的邻域结构,结合用于提取样本图像本身特征的第一特征提取网络,从两个角度来对样本图像的图像特征进行处理,考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,这样提取到的特征能够更加准确地体现出图像所属的类别,基于训练后的特征提取网络进行图像检索也能够得到更加准确的检索结果。



1. 一种特征提取网络训练方法,其特征在于,包括:

基于第一特征提取网络对样本图像进行处理,得到所述样本图像的第一特征,所述第一特征用于表征所述样本图像的内容特征;

基于第二特征提取网络对所述样本图像进行处理,得到所述样本图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述样本图像所属的类别;

基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,所述第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,所述预测相似度关系用于表征两个样本图像之间是否属于同一类别;

基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练。

2. 根据权利要求1所述的特征提取网络训练方法,其特征在于,所述基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,包括:

基于所述样本图像以及所述第一样本图像集合中第一样本图像的第一特征,确定所述样本图像与所述第一样本图像之间的第一相似度,所述第一样本图像为所述样本图像集合中与所述样本图像属于相同类别的图像;

基于所述样本图像以及与所述第一样本图像集合中第二样本图像的第一特征,确定所述样本图像与所述第二样本图像之间的第二相似度,第二样本图像为所述第一样本图像集合中与所述样本图像属于不同类别的图像;

基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,所述第二样本图像集合包括所述第一样本图像和所述第二样本图像;

基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;

基于所述第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系;

所述基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练,包括:

基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练;

基于所述第二预测相似度关系与目标相似度关系之间的差异,对所述第二特征提取网络进行训练。

3. 根据权利要求2所述的特征提取网络训练方法,其特征在于,所述基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,包括:

根据所述第一相似度和所述第二相似度,从所述第二样本图像集合中,确定所述样本图像对应的第三样本图像集合,所述第三样本图像集合包括所述第一相似度或所述第二相

似度满足条件的图像；

以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征和所述第三样本图像集合中样本图像的第一特征之间的距离；

对所述距离进行归一化,得到所述样本图像和所述第三样本图像集合中样本图像之间的第三相似度。

4. 根据权利要求3所述的特征提取网络训练方法,其特征在于,所述基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系,包括:

基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;

基于所述样本图像与所述第三样本图像集合中样本图像之间的第三相似度,对所述第三样本图像集合中样本图像进行排序,得到预测排序结果;

所述基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练,包括:

基于所述样本图像与所述第三样本图像集合中样本图像之间的第一相似度,对所述第三样本图像集合中样本图像进行排序,得到目标排序结果;

基于所述预测排序结果与目标排序结果,确定第二损失值,所述第二损失值用于指示所述排序结果和所述目标排序结果之间的差异;

基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

5. 一种图像处理方法,其特征在于,包括:

对图像进行处理,得到所述图像的第一特征,所述第一特征用于表征所述图像的内容特征;

对所述图像进行处理,得到所述图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述图像所属的类别;

基于所述图像的所述第一特征和所述第二特征,确定所述图像所匹配的目标图像;

基于所述目标图像进行图像推送。

6. 一种特征提取网络训练装置,其特征在于,包括:

处理单元,被配置为执行基于第一特征提取网络对样本图像进行处理,得到所述样本图像的第一特征,所述第一特征用于表征所述样本图像的内容特征;

所述处理单元,还被配置为执行基于第二特征提取网络对所述样本图像进行处理,得到所述样本图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述样本图像所属的类别;

确定单元,被配置为执行基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,所述第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,所述预测相似度关系用于表征两个样本图像之间是否属于同一类别;

训练单元,被配置为执行基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练。

7. 一种图像处理装置,其特征在于,包括:

处理单元,被配置为执行对所述图像进行处理,得到所述图像的第一特征,所述第一特征用于表征所述图像的内容特征;

所述处理单元,还被配置为执行对所述图像进行处理,得到所述图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述图像所属的类别;

确定单元,被配置为执行基于所述图像的所述第一特征和所述第二特征,确定所述图像所匹配的目标图像;

推送单元,被配置为执行基于所述目标图像进行图像推送。

8. 一种电子设备,其特征在于,包括:

一个或多个处理器;

用于存储所述一个或多个处理器可执行指令的一个或多个存储器;

其中,所述一个或多个处理器被配置为执行所述指令,以实现如权利要求1至7中任一项所述的特征提取网络训练方法,或实现如权利要求8至11任一项所述的图像处理方法。

9. 一种计算机可读存储介质,当所述计算机可读存储介质中的指令由电子设备的处理器执行时,使得所述电子设备能够执行如权利要求1至4中任一项所述的特征提取网络训练方法,或实现如权利要求5所述的图像处理方法。

10. 一种计算机程序产品,包括计算机指令,其特征在于,所述计算机指令被处理器执行时实现如权利要求1至4中任一项所述的特征提取网络训练方法,或实现如权利要求5所述的图像处理方法。

特征提取网络训练方法、图像处理方法及装置

技术领域

[0001] 本公开涉及图像处理领域,尤其涉及特征提取网络训练方法、图像处理方法及装置。

背景技术

[0002] 在图像处理领域,能够通过网络对两个图像进行处理,以确定出两个图像的之间的相似度,以此来确定两个图像是否相关,以此来作为图像检索的依据。

[0003] 目前,特征提取网络训练方法的关注点通常在于图像的嵌入表示的学习,训练时特征提取网络能够对样本图像的图像特征进行处理得到嵌入表示,然后基于嵌入表示对特征提取网络进行训练,以使得特征提取网络训练后能够准确提取图像的嵌入表示。

[0004] 上述基于嵌入表示训练网络的方式有个基本前提:假设样本图像周围的邻域结构非常均匀。但是其实样本图像的邻域结构一般不像假设的那么均匀,因而上述方式确定的嵌入表示并没有那么准确,自然地,该特征提取网络确定出图像之间的相似度也就不够准确,该特征提取网络用于图像检索时得到的检索结果不够准确。

发明内容

[0005] 本公开提供一种特征提取网络训练方法、图像处理方法及装置,能够提高特征提取网络的准确性。本公开的技术方案如下:

[0006] 根据本公开实施例的第一方面,提供一种特征提取网络训练方法,包括:

[0007] 基于第一特征提取网络对样本图像进行处理,得到所述样本图像的第一特征,所述第一特征用于表征所述样本图像的内容特征;

[0008] 基于第二特征提取网络对所述样本图像进行处理,得到所述样本图像的第二特征,第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述样本图像所属的类别;

[0009] 基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,所述第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,所述预测相似度关系用于表征两个样本图像之间是否属于同一类别;

[0010] 基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练。

[0011] 在一些实施例中,所述基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,包括:

[0012] 基于所述样本图像以及所述第一样本图像集合中第一样本图像的第一特征,确定所述样本图像与所述第一样本图像之间的第一相似度,所述第一样本图像为所述样本图像集合中与所述样本图像属于相同类别的图像;

[0013] 基于所述样本图像以及与所述第一样本图像集合中第二样本图像的第一特征,确

定所述样本图像与所述第二样本图像之间的第二相似度,第二样本图像为所述第一样本图像集合中与所述样本图像属于不同类别的图像;

[0014] 基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,所述第二样本图像集合包括所述第一样本图像和所述第二样本图像;

[0015] 基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;

[0016] 基于所述第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系;

[0017] 所述基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练,包括:

[0018] 基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练;

[0019] 基于所述第二预测相似度关系与目标相似度关系之间的差异,对所述第二特征提取网络进行训练。

[0020] 在一些实施例中,所述基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,包括:

[0021] 根据所述第一相似度和所述第二相似度,从所述第二样本图像集合中,确定所述样本图像对应的第三样本图像集合,所述第三样本图像集合包括所述第一相似度或所述第二相似度满足条件的图像;

[0022] 以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征和所述第三样本图像集合中样本图像的第一特征之间的距离;

[0023] 对所述距离进行归一化,得到所述样本图像和所述第三样本图像集合中样本图像之间的第三相似度。

[0024] 在一些实施例中,所述基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系,包括:

[0025] 基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;

[0026] 基于所述样本图像与所述第三样本图像集合中样本图像之间的第三相似度,对所述第三样本图像集合中样本图像进行排序,得到预测排序结果;

[0027] 所述基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练,包括:

[0028] 基于所述样本图像与所述第三样本图像集合中样本图像之间的第一相似度,对所述第三样本图像集合中样本图像进行排序,得到目标排序结果;

[0029] 基于所述预测排序结果与目标排序结果,确定第二损失值,所述第二损失值用于指示所述排序结果和所述目标排序结果之间的差异;

[0030] 基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0031] 在一些实施例中,所述基于所述样本图像的第二特征、所述样本图像与所述第二

样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,包括:

[0032] 以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征与所述第二样本图像集合中样本图像的第一特征之间的距离,所述距离为所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度。

[0033] 在一些实施例中,所述基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练,包括:

[0034] 基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;

[0035] 根据所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度,确定所述第二样本图像集合中样本图像中所述第一样本图像的第一中心,以及所述第二样本图像的第二中心;

[0036] 根据所述样本图像和所述第二样本图像集合中样本图像分别与所述第一中心以及所述第二中心之间的距离,获取第二损失值;

[0037] 基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0038] 在一些实施例中,所述第二特征提取网络包括两个全连接层和激活层;

[0039] 所述基于第二特征提取网络对所述样本图像进行处理,得到所述样本图像的第二特征,包括:

[0040] 基于所述第二特征提取网络中的所述两个全连接层,对所述样本图像进行前向计算,由所述第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出所述样本图像的第二特征。

[0041] 根据本公开实施例的第二方面,提供一种图像处理方法,包括:

[0042] 对所述图像进行处理,得到所述图像的第一特征,所述第一特征用于表征所述图像的内容特征;

[0043] 对所述图像进行处理,得到所述图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述图像所属的类别;

[0044] 基于所述图像的所述第一特征和所述第二特征,确定所述图像所匹配的目标图像;

[0045] 基于所述目标图像进行图像推送。

[0046] 在一些实施例中,所述对所述图像进行处理,得到所述图像的第一特征,包括:

[0047] 基于第一特征提取网络对所述图像进行处理,得到所述第一特征;

[0048] 所述对所述图像进行处理,得到所述图像的第二特征,包括:

[0049] 基于第二特征提取网络,对所述图像进行处理,得到所述图像的第二特征;

[0050] 其中,所述第一特征提取网络和所述第二特征提取网络分别基于第一样本图像集合中样本图像之间的预测相似度关系与目标预测相似度关系之间的差异训练得到,所述预测相似度关系基于样本图像的第一特征、第二特征以及所述样本图像所属类别确定,且所述预测相似度关系用于表征两个样本图像之间是否属于同一类别。

[0051] 在一些实施例中,所述基于所述图像的所述第一特征和所述第二特征,确定所述

图像所匹配的目标图像,包括:

[0052] 基于所述图像的第一特征与候选图像集合中候选图像的第一特征之间的第四相似度,从所述候选图像集合中,确定所述图像所匹配的目标图像,所述目标图像的所述第四相似度符合条件;

[0053] 基于所述图像的第二特征与所述目标图像的第二特征之间的第五相似度,对所述目标图像进行排序,得到所述目标图像的排序结果;

[0054] 所述基于所述目标图像进行图像推送,包括:

[0055] 按照所述排序结果对所述目标图像进行图像推送。

[0056] 在一些实施例中,所述第二特征提取网络包括两个全连接层和激活层;

[0057] 所述对所述图像进行处理,得到所述图像的第二特征,包括:

[0058] 基于第二特征提取网络中的两个全连接层,对所述图像进行前向计算,由所述第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出所述图像的第二特征。

[0059] 根据本公开实施例的第三方面,提供一种特征提取网络训练装置,包括:

[0060] 处理单元,被配置为执行基于第一特征提取网络对样本图像进行处理,得到所述样本图像的第一特征,所述第一特征用于表征所述样本图像的内容特征;

[0061] 所述处理单元,还被配置为执行基于第二特征提取网络对所述样本图像进行处理,得到所述样本图像的第二特征,第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述样本图像所属的类别;

[0062] 确定单元,被配置为执行基于所述样本图像的第一特征、第二特征以及所述样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,所述第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,所述预测相似度关系用于表征两个样本图像之间是否属于同一类别;

[0063] 训练单元,被配置为执行基于所述预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络和所述第二特征提取网络进行训练。

[0064] 在一些实施例中,所述确定单元被配置为执行:

[0065] 基于所述样本图像以及所述第一样本图像集合中第一样本图像的第一特征,确定所述样本图像与所述第一样本图像之间的第一相似度,所述第一样本图像为所述样本图像集合中与所述样本图像属于相同类别的图像;

[0066] 基于所述样本图像以及与所述第一样本图像集合中第二样本图像的第一特征,确定所述样本图像与所述第二样本图像之间的第二相似度,第二样本图像为所述第一样本图像集合中与所述样本图像属于不同类别的图像;

[0067] 基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,所述第二样本图像集合包括所述第一样本图像和所述第二样本图像;

[0068] 基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;

[0069] 基于所述第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系;

[0070] 所述训练单元被配置为执行：

[0071] 基于所述第一预测相似度关系与目标相似度关系之间的差异，对所述第一特征提取网络进行训练；

[0072] 基于所述第二预测相似度关系与目标相似度关系之间的差异，对所述第二特征提取网络进行训练。

[0073] 在一些实施例中，所述确定单元被配置为执行：

[0074] 根据所述第一相似度和所述第二相似度，从所述第二样本图像集合中，确定所述样本图像对应的第三样本图像集合，所述第三样本图像集合包括所述第一相似度或所述第二相似度满足条件的图像；

[0075] 以所述样本图像的第二特征作为系数，获取所述样本图像的第一特征和所述第三样本图像集合中样本图像的第一特征之间的距离；

[0076] 对所述距离进行归一化，得到所述样本图像和所述第三样本图像集合中样本图像之间的第三相似度。

[0077] 在一些实施例中，所述确定单元被配置为执行：

[0078] 基于所述第一相似度和所述第二相似度，确定第一损失值，所述第一损失值与所述第一相似度正相关，所述第一损失值与所述第二相似度负相关；

[0079] 基于所述样本图像与所述第三样本图像集合中样本图像之间的第三相似度，对所述第三样本图像集合中样本图像进行排序，得到预测排序结果；

[0080] 所述训练单元被配置为执行：基于所述样本图像与所述第三样本图像集合中样本图像之间的第一相似度，对所述第三样本图像集合中样本图像进行排序，得到目标排序结果；

[0081] 基于所述预测排序结果与目标排序结果，确定第二损失值，所述第二损失值用于指示所述排序结果和所述目标排序结果之间的差异；

[0082] 基于所述第一损失值和所述第二损失值，对所述第一特征提取网络进行训练。

[0083] 在一些实施例中，所述确定单元被配置为执行：

[0084] 以所述样本图像的第二特征作为系数，获取所述样本图像的第一特征与所述第二样本图像集合中样本图像的第一特征之间的距离，所述距离为所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度。

[0085] 在一些实施例中，所述确定单元和所述训练单元被配置为执行：

[0086] 基于所述第一相似度和所述第二相似度，确定第一损失值，所述第一损失值与所述第一相似度正相关，所述第一损失值与所述第二相似度负相关；

[0087] 根据所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度，确定所述第二样本图像集合中样本图像中所述第一样本图像的第一中心，以及所述第二样本图像的第二中心；

[0088] 根据所述样本图像和所述第二样本图像集合中样本图像分别与所述第一中心以及所述第二中心之间的距离，获取第二损失值；

[0089] 基于所述第一损失值和所述第二损失值，对所述第一特征提取网络进行训练。

[0090] 在一些实施例中，所述第二特征提取网络包括两个全连接层和激活层；

[0091] 所述处理单元被配置为执行基于所述第二特征提取网络中的所述两个全连接层，

对所述样本图像进行前向计算,由所述第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出所述样本图像的第二特征。

[0092] 根据本公开实施例的第四方面,提供一种图像处理装置,包括:

[0093] 处理单元,被配置为执行对所述图像进行处理,得到所述图像的第一特征,所述第一特征用于表征所述图像的内容特征;

[0094] 所述处理单元,还被配置为执行对所述图像进行处理,得到所述图像的第二特征,所述第二特征用于表征目标类别图像的内容特征分布状况,所述目标类别为所述图像所属的类别;

[0095] 确定单元,被配置为执行基于所述图像的所述第一特征和所述第二特征,确定所述图像所匹配的目标图像;

[0096] 推送单元,被配置为执行基于所述目标图像进行图像推送。

[0097] 在一些实施例中,所述处理单元被配置为执行:

[0098] 基于第一特征提取网络对所述图像进行处理,得到所述第一特征;

[0099] 所述处理单元被配置为执行:

[0100] 基于第二特征提取网络,对所述图像进行处理,得到所述第二特征;

[0101] 其中,所述第一特征提取网络和所述第二特征提取网络分别基于第一样本图像集合中样本图像之间的预测相似度关系与目标预测相似度关系之间的差异训练得到,所述预测相似度关系基于样本图像的第一特征、第二特征以及所述样本图像所属类别确定,且所述预测相似度关系用于表征两个样本图像之间是否属于同一类别。

[0102] 在一些实施例中,所述确定单元被配置为执行:

[0103] 基于所述图像的第一特征与候选图像集合中候选图像的第一特征之间的第四相似度,从所述候选图像集合中,确定所述图像所匹配的目标图像,所述目标图像的所述第四相似度符合条件;

[0104] 基于所述图像的第二特征与所述目标图像的第二特征之间的第五相似度,对所述目标图像进行排序,得到所述目标图像的排序结果;

[0105] 所述推送单元被配置为执行:

[0106] 按照所述排序结果对所述目标图像进行图像推送。

[0107] 在一些实施例中,所述第二特征提取网络包括两个全连接层和激活层;

[0108] 所述处理单元被配置为执行:

[0109] 基于第二特征提取网络中的两个全连接层,对所述图像进行前向计算,由所述第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出所述图像的第二特征。

[0110] 根据本公开实施例的第五方面,提供一种电子设备,电子设备包括:

[0111] 一个或多个处理器;

[0112] 用于存储所述一个或多个处理器可执行指令的一个或多个存储器;

[0113] 其中,所述一个或多个处理器被配置为执行所述指令,以实现第一方面任一项所述的特征提取网络训练方法或第二方面所述的图像处理方法。

[0114] 根据本公开实施例的第六方面,提供一种计算机可读存储介质,当所述计算机可读存储介质中的指令由电子设备的处理器执行时,使得所述电子设备能够执行第一方面任

一项所述的特征提取网络训练方法或第二方面所述的图像处理方法。

[0115] 根据本公开实施例的第七方面,提供一种计算机程序产品,包括一条或多条计算机指令,所述一条或多条计算机指令存储在计算机可读存储介质中。电子设备的一个或多个处理器能够从计算机可读存储介质中读取所述一条或多条计算机指令,所述一个或多个处理器执行所述一条或多条计算机指令,使得电子设备能够执行上述第一方面任一项所述的特征提取网络训练方法或第二方面所述的图像处理方法。

[0116] 本公开的实施例提供的技术方案至少带来以下有益效果:

[0117] 本公开实施例,新增了一种用于提取第二特征的第二特征提取网络,第二特征用于表征目标类别图像的内容特征分布状况,也即是该样本图像的邻域结构,结合用于提取样本图像本身特征的第一特征提取网络,从两个角度来对样本图像的图像特征进行处理,考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,这样提取到的特征能够更加准确地体现出图像所属的类别,基于训练后的特征提取网络进行图像检索也能够得到更加准确的检索结果。

[0118] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

附图说明

[0119] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理,并不构成对本公开的不当限定。

[0120] 图1是本申请实施例提供的一种特征提取网络训练方法和图像处理方法的实施环境的示意图。

[0121] 图2是根据一示例性实施例示出的一种特征提取网络训练方法的流程图。

[0122] 图3是根据一示例性实施例示出的一种特征提取网络训练方法的流程图。

[0123] 图4是根据一示例性实施例示出的一种特征提取网络训练方法的流程图。

[0124] 图5是根据一示例性实施例示出的一种图像的邻域结构的示意图。

[0125] 图6是根据一示例性实施例示出的一种特征提取网络训练方法的流程图。

[0126] 图7是根据一示例性实施例示出的一种特征提取网络训练方法的流程图。

[0127] 图8是根据一示例性实施例示出的一种样本图像的距离分布的示意图。

[0128] 图9是根据一示例性实施例示出的一种图像处理方法的流程图。

[0129] 图10是根据一示例性实施例示出的特征提取网络训练方法在CUB-200-2011数据集上的表现示意图。

[0130] 图11是根据一示例性实施例示出的一种特征提取网络训练装置框图。

[0131] 图12是根据一示例性实施例示出的一种特征提取网络训练装置框图。

[0132] 图13是根据一示例性实施例示出的一种电子设备的结构示意图。

[0133] 图14是根据一示例性实施例示出的一种终端的结构框图。

[0134] 图15是根据一示例性实施例示出的一种服务器的结构示意图。

具体实施方式

[0135] 为了使本领域普通人员更好地理解本公开的技术方案,下面将结合附图,对本公开实施例中的技术方案进行清楚、完整地描述。

[0136] 需要说明的是,本公开的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本公开的实施例能够以除了在这里图示或描述的那些以外的顺序实施。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0137] 本公开所涉及的用户信息可以为经用户授权或者经过各方充分授权的信息。

[0138] 下面针对本公开实施例涉及到的名词进行解释说明。

[0139] 在拓扑学和相关的数学领域中,邻域是拓扑空间中的基本概念。直觉上说,一个点的邻域是包含这个点的集合。

[0140] 拓扑流形的数学定义可以表述为:设 M 是豪斯多夫空间,若对任意一点 $x \in M$,都有 x 在 M 中的一个邻域 U 同胚于 m 维欧几里得空间 R^m 的一个开集,就称 M 是一个 m 维流形或 m 维拓扑流形。

[0141] 相关技术中,都是假设样本图像周围的邻域结构非常均匀,但是样本图像的邻域结构一般不像假设的那么均匀,以下通过两个示例来对邻域结构不均匀进行详细解释说明。

[0142] 示例一:在对抗学习中,对抗样本在某些输入维度上的小扰动可能会导致输出空间的较大偏差,而在其他维度上的扰动则不会,在其他维度上的扰动不会影响输出空间的原因是神经网络具有一定的鲁棒性。通过对抗样本在不同输入维度上的扰动的不同表现,正可以证明嵌入的邻域结构不像假设的那么均匀。

[0143] 示例二:如图1所示,包括两种图像,每种图像的图像数量为多个。对于其中一种图像,每个图像中包括猫,这些图像的类别为猫。对于另一种图像,每个图像中包括虎,这些图像的类别为虎。不同图像里包括的猫在外观或者品种上可能不同。对于类别为猫的一种图像,由于猫的纹理图案非常不同,因此这些图像在嵌入空间中其邻居可以在“纹理”维度上分布得很宽。相比之下,对于老虎来说,因为所有种类的老虎都具有相似的纹理,因而类别为虎的图像的邻居应该分布在狭窄的范围内。由图可知,在“纹理”维度上,猫和老虎应具有不同的分布形状,不同类别的样本图像应有自适应的邻域结构。

[0144] 通过上述两个示例能够充分说明样本图像的邻域结构其实并不均匀,正是考虑到这一点,本申请实施例中新增了一种第二特征提取网络,该第二特征提取网络能够提取到表示邻域结构的第二特征,以此来辅助第一特征提取网络训练,从而能够使得特征提取网络提取到更加准确的特征。具体内容可以参见下述实施例,在此先不作过多介绍。

[0145] 图2是本申请实施例提供的一种特征提取网络训练方法和图像处理方法的实施环境的示意图。该实施环境包括终端101,或者该实施环境包括终端101和图像处理平台102。终端101通过无线网络或有线网络与图像处理平台102相连。

[0146] 终端101是智能手机、游戏主机、台式计算机、平板电脑、电子书阅读器、MP3 (Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)

播放器或MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器,膝上型便携计算机,智能机器人,自助支付设备中的至少一种。终端101安装和运行有支持图像处理的应用程序,例如,该应用程序是系统应用、即时通讯应用、新闻推送应用、购物应用、在线视频应用、社交应用。

[0147] 示例性地,该终端101具有图像采集功能和图像处理功能,对采集到的图像进行处理,并根据处理结果执行相应的功能。该终端101独立完成该工作,也通过图像处理平台102为其提供数据服务,本申请实施例对此不作限定。对于特征提取网络的训练,可以由终端101执行,也可以由图像处理平台102执行,或者由终端101和图像处理平台102协作执行,本申请实施例对此不作限定。

[0148] 图像处理平台102包括一台服务器、多台服务器、云计算平台和虚拟化中心中的至少一种。图像处理平台102用于为支持图像处理的应用程序提供后台服务。可选地,图像处理平台102承担主要处理工作,终端101承担次要处理工作;或者,图像处理平台102承担次要处理工作,终端101承担主要处理工作;或者,图像处理平台102或终端101分别单独承担处理工作。或者,图像处理平台102和终端101两者之间采用分布式计算架构进行协同计算。

[0149] 可选地,该图像处理平台102包括至少一台服务器1021以及数据库1022,该数据库1022用于存储数据,在本申请实施例中,该数据库1022中存储有样本图像,为至少一台服务器1021提供数据服务。

[0150] 服务器是独立的物理服务器,也是多个物理服务器构成的服务器集群或者分布式系统,还是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。终端是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等,但并不局限于此。

[0151] 本领域技术人员知晓,上述终端101、服务器1021的数量更多或更少。比如上述终端101、服务器1021仅为一个,或者上述终端101、服务器1021为几十个或几百个,或者更多数量,本申请实施例对终端或服务器的数量和设备类型不加以限定。

[0152] 图3是根据一示例性实施例示出的一种特征提取网络训练方法的流程图,如图3所示,特征提取网络训练方法可以由电子设备执行,包括以下步骤。

[0153] 在步骤S21中,基于第一特征提取网络对样本图像进行处理,得到该样本图像的第一特征,该第一特征用于表征该样本图像的内容特征。

[0154] 在步骤S22中,基于第二特征提取网络对该样本图像进行处理,得到该样本图像的第二特征,第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该样本图像所属的类别。

[0155] 在步骤S13中,基于该样本图像的第一特征、第二特征以及该样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,该第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,该预测相似度关系用于表征两个样本图像之间是否属于同一类别。

[0156] 在步骤S24中,基于该预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络和该第二特征提取网络进行训练。

[0157] 本公开实施例,新增了一种用于提取第二特征的第二特征提取网络,第二特征用于表征目标类别图像的内容特征分布状况,也即是该样本图像的邻域结构,结合用于提取

样本图像本身特征的第一特征提取网络,从两个角度来对样本图像的图像特征进行处理,考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,这样提取到的特征能够更加准确地体现出图像所属的类别,基于训练后的特征提取网络进行图像检索也能够得到更加准确的检索结果。

[0158] 在一些实施例中,该基于该样本图像的第一特征、第二特征以及该样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,包括:

[0159] 基于该样本图像以及该第一样本图像集合中第一样本图像的第一特征,确定该样本图像与该第一样本图像之间的第一相似度,该第一样本图像为该样本图像集合中与该样本图像属于相同类别的图像;

[0160] 基于该样本图像以及与该第一样本图像集合中第二样本图像的第一特征,确定该样本图像与该第二样本图像之间的第二相似度,第二样本图像为该第一样本图像集合中与该样本图像属于不同类别的图像;

[0161] 基于该样本图像的第二特征、该样本图像与该第二样本图像集合中样本图像的第一特征,确定该样本图像和该第二样本图像集合中样本图像之间的第三相似度,该第二样本图像集合包括该第一样本图像和该第二样本图像;

[0162] 基于该第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;

[0163] 基于该第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系;

[0164] 该基于该预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络 and 该第二特征提取网络进行训练,包括:

[0165] 基于该第一预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络进行训练;

[0166] 基于该第二预测相似度关系与目标相似度关系之间的差异,对该第二特征提取网络进行训练。

[0167] 在一些实施例中,该基于该样本图像的第二特征、该样本图像与该第二样本图像集合中样本图像的第一特征,确定该样本图像和该第二样本图像集合中样本图像之间的第三相似度,包括:

[0168] 根据该第一相似度和该第二相似度,从该第二样本图像集合中,确定该样本图像对应的第三样本图像集合,该第三样本图像集合包括该第一相似度或该第二相似度满足条件的图像;

[0169] 以该样本图像的第二特征作为系数,获取该样本图像的第一特征和该第三样本图像集合中样本图像的第一特征之间的距离;

[0170] 对该距离进行归一化,得到该样本图像和该第三样本图像集合中样本图像之间的第三相似度。

[0171] 在一些实施例中,该基于该第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系,包括:

[0172] 基于该第一相似度和该第二相似度,确定第一损失值,该第一损失值与该第一相

似度正相关,该第一损失值与该第二相似度负相关;

[0173] 基于该样本图像与该第三样本图像集合中样本图像之间的第三相似度,对该第三样本图像集合中样本图像进行排序,得到预测排序结果;

[0174] 该基于该第一预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络进行训练,包括:

[0175] 基于该样本图像与该第三样本图像集合中样本图像之间的第一相似度,对该第三样本图像集合中样本图像进行排序,得到目标排序结果;

[0176] 基于该预测排序结果与目标排序结果,确定第二损失值,该第二损失值用于指示该排序结果和该目标排序结果之间的差异;

[0177] 基于该第一损失值和该第二损失值,对该第一特征提取网络进行训练。

[0178] 在一些实施例中,该基于该样本图像的第二特征、该样本图像与该第二样本图像集合中样本图像的第一特征,确定该样本图像和该第二样本图像集合中样本图像之间的第三相似度,包括:

[0179] 以该样本图像的第二特征作为系数,获取该样本图像的第一特征与该第二样本图像集合中样本图像的第一特征之间的距离,该距离为该样本图像与该第二样本图像集合中样本图像之间的第三相似度。

[0180] 在一些实施例中,该基于该第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;基于该第一预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络进行训练,包括:

[0181] 基于该第一相似度和该第二相似度,确定第一损失值,该第一损失值与该第一相似度正相关,该第一损失值与该第二相似度负相关;

[0182] 根据该样本图像与该第二样本图像集合中样本图像之间的第三相似度,确定该第二样本图像集合中样本图像中该第一样本图像的第一中心,以及该第二样本图像的第二中心;

[0183] 根据该样本图像和该第二样本图像集合中样本图像分别与该第一中心以及该第二中心之间的距离,获取第二损失值;

[0184] 基于该第一损失值和该第二损失值,对该第一特征提取网络进行训练。

[0185] 在一些实施例中,该第二特征提取网络包括两个全连接层和激活层;

[0186] 该基于第二特征提取网络对该样本图像进行处理,得到该样本图像的第二特征,包括:

[0187] 基于该第二特征提取网络中的该两个全连接层,对该样本图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该样本图像的第二特征。

[0188] 图4是根据一示例性实施例示出的一种特征提取网络训练方法的流程图,如图4所示,特征提取网络训练方法可以由电子设备执行,包括以下步骤。

[0189] 在步骤S31中,对该图像进行处理,得到该图像的第一特征,该第一特征用于表征该图像的内容特征。

[0190] 在步骤S32中,对该图像进行处理,得到该图像的第二特征,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该图像所属的类别。

[0191] 在步骤S33中,基于该图像的该第一特征和该第二特征,确定该图像所匹配的目标图像。

[0192] 在步骤S34中,基于该目标图像进行图像推送。

[0193] 本公开实施例在进行图像检索时,不仅利用到了直接通过图像特征得到了第一特征,该第一特征用于表示图像本身内容的特征,还利用到了用于表示该图像的邻域结构的第二特征,结合两种特征,从两个角度来对对图像进行描述,以此匹配到的目标图像与图像在图像内容方面和所属类别方面等均能够更加贴近,以此进行图像推送更加准确,效果更好。

[0194] 在一些实施例中,所述对所述图像进行处理,得到所述图像的第一特征,包括:

[0195] 基于第一特征提取网络对所述图像进行处理,得到所述第一特征;

[0196] 所述对所述图像进行处理,得到所述图像的第二特征,包括:

[0197] 基于第二特征提取网络,对所述图像进行处理,得到所述图像的第二特征;

[0198] 其中,所述第一特征提取网络和所述第二特征提取网络分别基于第一样本图像集合中样本图像之间的预测相似度关系与目标预测相似度关系之间的差异训练得到,所述预测相似度关系基于样本图像的第一特征、第二特征以及所述样本图像所属类别确定,且所述预测相似度关系用于表征两个样本图像之间是否属于同一类别。

[0199] 在一些实施例中,所述第一特征提取网络和所述第二特征提取网络的训练过程包括:

[0200] 基于所述样本图像以及所述第一样本图像集合中第一样本图像的第一特征,确定所述样本图像与所述第一样本图像之间的第一相似度,所述第一样本图像为所述样本图像集合中与所述样本图像属于相同类别的图像;

[0201] 基于所述样本图像以及与所述第一样本图像集合中第二样本图像的第一特征,确定所述样本图像与所述第二样本图像之间的第二相似度,第二样本图像为所述第一样本图像集合中与所述样本图像属于不同类别的图像;

[0202] 基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,所述第二样本图像集合包括所述第一样本图像和所述第二样本图像;

[0203] 基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系;

[0204] 基于所述第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系;

[0205] 基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练;

[0206] 基于所述第二预测相似度关系与目标相似度关系之间的差异,对所述第二特征提取网络进行训练。

[0207] 在一些实施例中,所述基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,包括:

[0208] 根据所述第一相似度和所述第二相似度,从所述第二样本图像集合中,确定所述

样本图像对应的第三样本图像集合,所述第三样本图像集合包括所述第一相似度或所述第二相似度满足条件的图像;

[0209] 以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征和所述第三样本图像集合中样本图像的第一特征之间的距离;

[0210] 对所述距离进行归一化,得到所述样本图像和所述第三样本图像集合中样本图像之间的第三相似度。

[0211] 在一些实施例中,所述基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系,包括:

[0212] 基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;

[0213] 基于所述样本图像与所述第三样本图像集合中样本图像之间的第三相似度,对所述第三样本图像集合中样本图像进行排序,得到预测排序结果;

[0214] 所述基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练,包括:

[0215] 基于所述样本图像与所述第三样本图像集合中样本图像之间的第一相似度,对所述第三样本图像集合中样本图像进行排序,得到目标排序结果;

[0216] 基于所述预测排序结果与目标排序结果,确定第二损失值,所述第二损失值用于指示所述排序结果和所述目标排序结果之间的差异;

[0217] 基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0218] 在一些实施例中,所述基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,包括:

[0219] 以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征与所述第二样本图像集合中样本图像的第一特征之间的距离,所述距离为所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度。

[0220] 在一些实施例中,所述第一特征提取网络和所述第二特征提取网络的训练过程包括:

[0221] 基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;

[0222] 根据所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度,确定所述第二样本图像集合中样本图像中所述第一样本图像的第一中心,以及所述第二样本图像的第二中心;

[0223] 根据所述样本图像和所述第二样本图像集合中样本图像分别与所述第一中心以及所述第二中心之间的距离,获取第二损失值;

[0224] 基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0225] 在一些实施例中,该基于该图像的该第一特征和该第二特征,确定该图像所匹配的目标图像,包括:

[0226] 基于该图像的第一特征与候选图像集合中候选图像的第一特征之间的第四相似度,从该候选图像集合中,确定该图像所匹配的目标图像,该目标图像的该第四相似度符合

条件;

[0227] 基于该图像的第二特征与该目标图像的第二特征之间的第五相似度,对该目标图像进行排序,得到该目标图像的排序结果;

[0228] 该基于该目标图像进行图像推送,包括:

[0229] 按照该排序结果对该目标图像进行图像推送。

[0230] 在一些实施例中,所述第二特征提取网络包括两个全连接层和激活层;

[0231] 该对该图像进行处理,得到该图像的第二特征,包括:

[0232] 基于第二特征提取网络中的两个全连接层,对该图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该图像的第二特征。

[0233] 图5是根据一示例性实施例示出的一种特征提取网络训练方法的流程图,如图5所示,该方法包括以下步骤。

[0234] 在步骤S41中,电子设备获取第一样本图像集合,该第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像。

[0235] 电子设备获取该第一样本图像集合,通过该第一样本图像集合中的多个样本图像对特征提取网络进行训练,使得该特征提取网络能准确提取到对图像的特征,以便于后续能够基于提取到的特征进行图像检索等。

[0236] 该第一样本图像集合可以存储于不同位置,相应地,该电子设备可以通过不同的方式获取该样本图像集合。

[0237] 在一些实施例中,该第一样本图像集合可以存储于图像数据库中,相应地,电子设备可以从图像数据库中提取该第一样本图像集合。例如,该图像数据库可以为语料库。

[0238] 在另一些实施例中,该第一样本图像集合也可以存储于该电子设备中,相应地,电子设备可以从本地存储中提取该第一样本图像集合。

[0239] 本申请对该第一样本图像集合的存储位置以及获取方式不作具体限定。

[0240] 在步骤S42中,电子设备对第一样本图像集合中样本图像进行特征提取,得到样本图像的图像特征。

[0241] 电子设备获取到样本图像后,可以对样本图像进行特征提取,得到样本图像的图像特征。图像特征能够表征图像中内容的特点。

[0242] 具体的,电子设备获取到第一样本图像集合后,该第一样本图像集合中包括多个样本图像。电子设备可以对第一样本图像集合中每个样本图像进行特征提取,得到每个样本图像的图像特征。

[0243] 对于特征提取,电子设备可以通过多种方式进行特征提取。

[0244] 在一些实施例中,电子设备可以通过特征提取网络进行特征提取。在一些实施例中,该特征提取网络可以包括一个或多个卷积层,该一个或多个卷积层用于对输入的图像进行卷积处理。相应地,该特征提取网络能够对输入的图像进行卷积处理,得到该图像的图像特征。

[0245] 对于该特征提取网络,该特征提取网络可以为CNN网络,也可以为其他网络,本申请实施例对此不作限定。在一个具体示例中,该特征提取网络可以为图像处理网络的骨干网络(Backbone),比如ResNet的骨干网络。

[0246] 该步骤S41和步骤S42为获取第一样本图像集合中样本图像的图像特征的过程,该

图像特征是指对样本图像的像素值进行处理得到的图像特征,通过该图像特征,能够较好地对该样本图像的图像内容进行表征。

[0247] 在步骤S43中,电子设备基于第一特征提取网络对该样本图像的图像特征进行处理,得到该样本图像的第一特征。

[0248] 电子设备获取到样本图像的图像特征后,还可以进一步对图像特征进行处理,以提取到更准确的特征。该第一特征用于表征该样本图像的内容特征,也即是指用于体现该样本图像本身的图像内容的特征。

[0249] 在该步骤S43中,通过第一特征提取网络对图像特征进行处理,得到该样本图像的图像内容方面的第一特征。

[0250] 在一些实施例中,通常用于表示图像内容的特点的特征称为嵌入特征或嵌入。样本图像的第一特征则可以称为样本嵌入 z ,该第一特征提取网络也可以称为样本嵌入网络(Sample Embedding Network)。假设上述步骤S42提取到的图像特征用 h 来表示。样本嵌入网络则用于提取样本嵌入 z 。

[0251] 对于第一特征提取网络,该第一特征提取网络包括一个或多个全连接层。通过第一特征提取网络可以对图像特征 h 进行卷积处理,以得到第一特征 z 。

[0252] 该步骤S41至S43是基于第一特征提取网络对样本图像进行处理,得到该样本图像的第一特征的一种可能实现方式,上述仅以先获取到样本图像的图像特征后,将图像特征输入至第一特征提取网络,由第一特征提取网络对图像特征进行处理,得到第一特征为例。在另一些实施例中,可以直接将样本图像输入第一特征提取网络中,由第一特征提取网络对其进行处理,得到第一特征。本公开实施例对具体采用哪种实现方式不作限定。

[0253] 在步骤S44中,电子设备基于第二特征提取网络对该样本图像的图像特征进行处理,得到该样本图像的第二特征,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该样本图像所属的类别。

[0254] 在对图像特征进行处理,以提取更准确的特征时,除了用于表示图像内容的第一特征之外,还可以对该样本图像的邻域结构进行分析,以此来为样本图像进行特征表示。

[0255] 可以理解地,对于图像内容相似的图像,这些图像会聚集在一起,因而对于一个图像,该图像的邻域是与该图像的图像内容相似的图像。

[0256] 例如,如图1所示,包括两种图像,每种图像的图像数量为多个。对于其中一种图像,每个图像中包括猫,这些图像的分类为猫。对于另一种图像,每个图像中包括虎,这些图像的分类为虎。不同图像里包括的猫在外观或者品种上可能不同。对于分类为猫的一种图像,由于猫的纹理图案非常不同,因此这些图像在嵌入空间中其邻居可以在“纹理”维度上分布得很宽。相比之下,对于老虎来说,因为所有种类的老虎都具有相似的纹理,因而分类为虎的图像的邻居应该分布在狭窄的范围内。由图可知,在“纹理”维度上,猫和老虎应具有不同的分布形状,不同类别的样本图像应有自适应的邻域结构。

[0257] 本公开即考虑到这一点,因而,通过分析与该样本图像属于同一类别的样本图像的分布,来为该样本图像提供第二特征,通过第二特征能够对样本图像在邻域结构上的特点进行表示。

[0258] 在一些实施例中,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该样本图像所属的类别。该第二特征用于表征目标类别图像的分布情况或者说目

标类别图像的拓扑结构,或者说也可以说,第二特征用于从图像分布的角度来描述该目标类别图像。也即是,第二特征用于表示样本图像的邻域结构。因而,也可以称第二特征为邻域结构特征或结构嵌入。第二特征提取网络为结构嵌入网络(Structure Embedding Network)。

[0259] 在一些实施例中,该第二特征提取网络包括两个全连接层和激活层。

[0260] 在本公开实施例中,该全连接层用于将该样本图像的全局特征来映射到邻域分布空间中,得到该样本图像的邻域结构表示。

[0261] 激活层用于将输入映射到输出,该激活层用于引入非线性因素,以逼近任何非线性函数。

[0262] 相应地,该步骤S44中,电子设备基于该第二特征提取网络中的该两个全连接层,对该样本图像的图像特征进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该样本图像的第二特征。

[0263] 通过两个全连接层,使得第二特征提取网络能够学习提取到嵌入的局部自适应邻域结构,得到用于表示该局部自适应邻域结构的第二特征,通过该第二特征对第一特征进行微调,能够得到样本图像更好的表征。

[0264] 上述步骤S41、S42和步骤S44是基于第二特征提取网络对该样本图像进行处理,得到该样本图像的第二特征的一种可能实现方式,上述仅以先获取到样本图像的图像特征后,将图像特征输入至第二特征提取网络,由第二特征提取网络对图像特征进行处理,得到第二特征为例。在另一些实施例中,可以直接将样本图像输入第二特征提取网络中,由第二特征提取网络对其进行处理,得到第二特征。

[0265] 相应的,在第二特征提取网络包括两个全连接层和激活层的实施例中,该步骤S44中,电子设备基于该第二特征提取网络中的该两个全连接层,对该样本图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该样本图像的第二特征。本公开实施例对具体采用哪种实现方式不作限定。

[0266] 在步骤S45中,电子设备基于该样本图像以及该第一样本图像集合中第一样本图像的第一特征,确定该样本图像与该第一样本图像之间的第一相似度,该第一样本图像为该样本图像集合中与该样本图像属于相同类别的图像。

[0267] 在得到样本图像的第一特征和第二特征后,电子设备可以进一步确定二者是否足够准确,以衡量第一特征提取网络和第二特征提取网络的网络参数是否合适。该步骤S45即为基于第一特征来分析第一特征提取网络的性能的过程。

[0268] 在确定样本图像的第一特征是否准确时,可以采用自监督学习的方式。在对样本图像进行分析时,可以认为该样本图像为锚样本。与该样本图像属于相同类别的图像为正样本,也即是第一样本图像为正样本。与该样本图像属于不同类别的图像为负样本,也即是第二样本图像为负样本。则样本图像和第一样本图像可以组成正样本对,样本图像和第二样本图像可以组成负样本对。该步骤S45则针对正样本对来进行分析。

[0269] 在一些实施例中,上述在确定样本图像的第一特征和第二特征时,也可以确定第一样本图像和第二样本图像的第一特征和第二特征。后续分析样本图像与第一样本图像或第二样本图像之间的相似度时,可以通过各自的第一特征或第二特征来实现。

[0270] 具体地,在确定该样本图像与该第一样本图像之间的第一相似度时,电子设备可

以获取样本图像的第一特征和第一样本图像的第一特征之间的相似度作为第一相似度。

[0271] 在步骤S46中,电子设备基于该样本图像以及与该第一样本图像集合中第二样本图像的第一特征,确定该样本图像与该第二样本图像之间的第二相似度,第二样本图像为该第一样本图像集合中与该样本图像属于不同类别的图像。

[0272] 在得到样本图像的第一特征和第二特征后,电子设备可以进一步确定二者是否足够准确,以衡量第一特征提取网络和第二特征提取网络的网络参数是否合适。该步骤S46即为基于第二特征来分析第二特征提取网络的性能的过程。

[0273] 在确定样本图像的第二特征是否准确时,可以采用自监督学习的方式。在对样本图像进行分析时,可以认为该样本图像为锚样本。与该样本图像属于不同类别的图像为负样本,也即是第二样本图像为负样本。样本图像和第二样本图像可以组成负样本对。该步骤S45则针对负样本对来进行分析。

[0274] 在一些实施例中,上述在确定样本图像的第一特征和第二特征时,也可以确定第一样本图像和第二样本图像的第一特征和第二特征。后续分析样本图像与第一样本图像或第二样本图像之间相似度时,可以通过各自的第一特征或第二特征来实现。

[0275] 具体地,在确定该样本图像与该第二样本图像之间的第二相似度时,电子设备可以获取样本图像的第一特征和第二样本图像的第一特征之间的相似度作为第二相似度。

[0276] 例如,在一个具体示例中,对于上述第一相似度和第二相似度,可以理解为两个图像之间的距离,具体可以通过下述公式一获得得到。

[0277] $dist(x_i, x_j) \triangleq d(z_i, z_j)$, 公式一

[0278] 其中,所有的样本图像可以通过 $X = \{x_i\}_{i=1}^n$ 表示, x_i, x_j 分别表示样本图像 x_i 和第一样本图像 x_j , 或者 x_i, x_j 分别表示样本图像 x_i 和第二样本图像 x_j 。 $d(\cdot, \cdot)$ 是嵌入空间中的距离度量函数,例如汉明距离,欧式距离和余弦距离。

[0279] 在步骤S47中,电子设备基于该样本图像的第二特征、该样本图像与该第二样本图像集合中样本图像的第一特征,确定该样本图像和该第二样本图像集合中样本图像之间的第三相似度,该第二样本图像集合包括该第一样本图像和该第二样本图像。

[0280] 通过上述步骤S45和步骤S46,电子设备通过第一特征,获取到了正样本对和负样本对的图像之间的相似度进行了分析。考虑到第二特征仅用于表示图像的邻域结构,在应用该第二特征确定图像之间相似度时,还需要借助第一特征。电子设备可以再综合第一特征和第二特征,获取正样本对和负样本对的图像之间的相似度。

[0281] 这样针对第一特征,获取到了一个相似度,又根据第一特征和第二特征,获取到了另一个相似度。针对第一特征则可以通过前一个相似度来确定第一特征提取网络的性能。针对第二特征,则可以通过后一个相似度来确定第二特征提取网络的性能。

[0282] 在一些实施例中,在确定第三相似度时,可以通过不同方式来实现。下面提供两种不同的方式,可以由相关技术人员根据图像处理需求自由选择采用哪种方式,本公开实施例对具体采用哪种方式不作限定。

[0283] 方式一、根据该第一相似度和该第二相似度,从该第二样本图像集合中,确定该样本图像对应的第三样本图像集合,该第三样本图像集合包括该第一相似度或该第二相似度满足条件的图像;以该样本图像的第二特征作为系数,获取该样本图像的第一特征和该第三样本图像集合中样本图像的第一特征之间的距离;对该距离进行归一化,得到该样本图

像和该第三样本图像集合中样本图像之间的第三相似度。

[0284] 在该方式一中,电子设备可以先根据第一相似度和第二相似度,对第一样本图像集合中所有除样本图像之外的样本图像进行排序。其中,第二样本图像集合即为第一样本图像集合中除了该样本图像之外的其他样本图像所组成的集合。

[0285] 可以理解地,如果是降序排列,则第二样本图像集合中越靠前的样本图像与样本图像越相似。如果是升序排列,则第二样本图像集合中越靠后的样本图像与样本图像越相似。在此,根据排序情况,从第二样本图像集合中筛选出与样本图像最相似的一些样本图像,在此将筛选出来的样本图像所组成的集合称为第三样本图像集合。后续再基于第三样本图像集合中样本图像进行训练。

[0286] 通常,具有高度相似性的样本对被视为高置信度监督信号。在这种情况下,可以认为与样本图像更相似的一些样本图像对邻域结构学习有利,而不太相似的那些样本图像则对邻域结构学习作用不大,甚至会影响学习结果,因而,通过基于相似度对第二样本图像集合中样本图像进行筛选,能够有效减少后续的训练计算量,提高训练效率,并能够保证训练结果足够准确。

[0287] 例如,在一个具体示例中,对于样本图像(锚样本),首先可以根据第一相似度和第二相似度,对第二样本图像集合中样本图像按照相似性降序排列,然后提取第二样本图像集合中前K个样本图像 n_1, \dots, n_k ,也即是第三样本图像集合中包含有该K个样本图像。对于第三样本图像集合中K个样本图像,电子设备可以基于第二特征 ω_i 和第一特征 z_i ,确定样本图像与第三样本图像集合中样本图像之间的第三相似度。

[0288] 在一些实施例中,该第三相似度可以理解为自适应距离 $d_{strc}(z_i, n_k)$ 。具体可以通过下述公式二实现:

$$[0289] \quad d_{strc}(z_i, n_k) = \sum_{d=1}^D \omega_i^{(d)} \cdot (z_i^{(d)} \oplus n_k^{(d)}) = \omega_i^T (z_i \oplus n_k), \quad \text{公式二}$$

[0290] 其中, ω_i^T 为第二特征 ω_i 的转置。 $d_{strc}(\cdot, \cdot)$ 是自适应距离函数, \oplus 是异或操作。

[0291] 假设上述通过相似性降序排列,则对于 $\forall p, q \in [1, K]$,满足下述公式三:

$$[0292] \quad \text{if } p < q, \text{ then } d_{strc}(z_i, n_p) < d_{strc}(z_i, n_q), \quad \text{公式三}$$

[0293] 在另一些实施例中,也可以不将上述自适应距离作为第三相似度,而是对其进行进一步映射处理,得到第三相似度。例如,可以将L2归一化后的第一特征的内积作为第三相似度,具体通过下述公式四,将自适应距离为自适应相似度。

$$[0294] \quad s_{strc}(z_i, n_k) = \frac{1}{D} \sum_{d=1}^D \omega_i^{(d)} - 2d_{strc}(z_i, n_k) = \omega_i^T (\tilde{z}_i^T \tilde{n}_k), \quad \text{公式四}$$

[0295] 其中, $s_{strc}(z_i, n_k)$ 为第三相似度, \tilde{z}_i 和 \tilde{n}_k 为归一化后的第一特征。

[0296] 方式二、以该样本图像的第二特征作为系数,获取该样本图像的第一特征与该第二样本图像集合中样本图像的第一特征之间的距离,该距离为该样本图像与该第二样本图像集合中样本图像之间的第三相似度。

[0297] 在该方式二中,电子设备并未根据第一相似度对第二样本图像集合中样本图像进行筛选,将第二样本图像集合中所有样本图像均作为训练依据,后续基于第二样本图像集合中所有样本图像均进行后续的损失值获取以及训练过程。具体确定相似度的过程与上述

方式一同理,只是不对第二样本图像集合中样本图像进行筛选,在此不多做赘述。

[0298] 省去了基于相似度进行筛选的步骤,能够减少因筛选而增加的计算量和耗时,将第二样本图像集合中所有样本图像均考虑在内,能够得到更加全面的相似度比较,后续训练时,能够使得网络学习到对更多种数据的处理,能够提高网络的处理性能。

[0299] 在步骤S48中,电子设备基于该第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系,基于该第一预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络进行训练。

[0300] 电子设备获取到第一相似度、第二相似度以及第三相似度后,该第一相似度基于属于相同类别的两个样本图像的第一特征确定,该第二相似度基于属于不同类别的两个样本图像的第一特征确定,该第三相似度基于该样本图像的第二特征、该样本图像与其他样本图像的第一特征确定。

[0301] 也即是,第一相似度和第二相似度为基于第一特征得到的,第三相似度基于第一特征和第二特征得到,这样通过第一相似度、第二相似度和第三相似度,对第一特征提取网络进行训练,不仅仅是基于第一特征进行训练,还考虑到了样本图像的邻域结构,这样考虑到了不同样本图像的邻域结构可能不均匀的情况,能够使得训练后第一特征提取网络在进行特征提取时也考虑到该邻域结构,从而得到更加准确的第一特征。

[0302] 在第三相似度确定方式不同时,基于得到的相似度对第一特征提取网络进行训练的方式也可以不同。

[0303] 对应于上述方式一,电子设备可以采用点排序学习的方式进行训练。具体地,可以通过下述步骤一至步骤五实现。

[0304] 步骤一、电子设备可以基于该第一相似度和该第二相似度,确定第一损失值,该第一损失值与该第一相似度正相关,该第一损失值与该第二相似度负相关。

[0305] 在训练过程中,可以通过该第一损失值,使得样本图像和第一样本图像之间的第一相似度越来越大,样本图像和第二样本图像之间的第二相似度越来越小,从而使得特征提取得到的第一特征能够明确将正负样本分开,且使得正样本之间越发靠近,负样本之间越发靠近。

[0306] 例如,在一个具体示例中,电子设备可以基于下述公式五获取第一损失值 L_1 。

[0307] $L_1 = d_{i,j} - d_{i,p} + \alpha$,公式五

[0308] 其中, L_1 为第一损失值, $d_{i,j}$ 为样本图像 x_i 与第一样本图像 x_j 之间的第一相似度, $d_{i,p}$ 为样本图像 x_i 与第二样本图像 x_p 之间的第二相似度, α 为目标间隔。该目标间隔为预设置的嵌入空间中正负样本之间的间隔。在此仅以该第一损失值的获取过程通过三元组损失函数来表示为例进行说明,还可以通过其他损失函数获取该第一损失值,比如,多相似度损失函数,本公开实施例对此不作限定。

[0309] 步骤二、电子设备基于该样本图像与每个该第三样本图像集合中样本图像之间的第三相似度,对该多个第三样本图像集合中样本图像进行排序,得到预测排序结果。

[0310] 该预测排序结果也即为一种第一预测相似度关系,在此将基于第一相似度排序的目标排序结果作为训练目标,然后通过步骤二中基于第三相似度排序得到的预测排序结果作为网络预测的结果,通过训练过程,使得预测排序结果接近目标排序结果,能够使得基于第三相似度排序的结果更加准确,这样可以使得特征提取网络提取到的第一特征和第二特

征均比较准确。

[0311] 步骤三、电子设备基于该样本图像与该第三样本图像集合中样本图像之间的第一相似度,对该第三样本图像集合中样本图像进行排序,得到目标排序结果。

[0312] 结构嵌入能够从嵌入空间中样本的局部流形中学习。由于流形可以完全通过其子区域之间的相对邻近度来表征,则距离之间的比较信息(例如,排序)就足以表征流形。在此则针对样本邻居的排序信息来指导结构嵌入的学习。

[0313] 步骤四、电子设备基于该预测排序结果与目标排序结果,确定第二损失值,该第二损失值用于指示该排序结果和该目标排序结果之间的差异。

[0314] 对于上述公式五,由于 $\sum \omega_i^{(d)}$,对于锚样本来说是常数,因此保留 $s_{strc}(z_i, n_k)$ 的顺序,等价于维持 $d_{strc}(z_i, n_k)$ 的顺序。在一个具体示例中,电子设备可以通过下述公式六,来确定第二损失值。

$$[0315] \quad \mathcal{L}_{strc} = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^{K-1} [s_{strc}(z_i, n_{k+1}) - s_{strc}(z_i, n_k) + \beta]_+, \quad \text{公式六}$$

[0316] 其中, \mathcal{L}_{strc} 为第二损失值。 β 为目标间隔。

[0317] 步骤五、电子设备基于该第一损失值和该第二损失值,对该第一特征提取网络进行训练。

[0318] 获取到第一损失值和第二损失值后,第一损失值可以体现第一特征的准确性,第二损失值可以体现第二特征的准确性,通过两个损失值对第一特征提取网络进行训练,能够使得第一特征提取网络考虑图像内容和邻域结构两个因素,提取到更加准确的第一特征。

[0319] 在一些实施例中,电子设备可以对第一损失值和第二损失值进行加权,得到目标损失值,基于目标损失值对第一特征提取网络进行训练。具体加权所采用的权重,可以由相关技术人员根据需求进行设置,例如,权重均为1,则该加权步骤也即是求和步骤。本公开实施例对此不作限定。

[0320] 在一个具体示例中,电子设备可以通过下述公式七,对第一损失值和第二损失值进行加权,得到目标损失值。

$$[0321] \quad \mathcal{L} = \mathcal{L}_{metric} + \lambda \mathcal{L}_{strc}, \quad \text{公式七}$$

[0322] 其中, λ 是平衡因子,也即是第二损失值的权重。

[0323] 例如,如图6所示,对于小批图像,可以输入骨干网络中,由骨干网络进行特征提取,得到图像特征后,分别输入到样本嵌入网络和结构嵌入网络中,样本嵌入网络能够对图像特征进行处理得到第一特征,结构嵌入网络能够对图像特征进行处理得到第二特征,然后基于第一特征和第二特征,可以分别获取第一损失值(度量学习损失)和第二损失值(结构学习损失),进而对两个网络进行训练。图6中输入的小批图像仅以方框示意,其具体用于表示图像。在此提供了图7,图7提供了一种输入图像的示意,输入的小批图像中针对所属类别组成不同的图像对输入骨干网络中。

[0324] 针对上述方式二,电子设备可以采用组排序学习进行训练。该组排序学习可以通过下述步骤一至步骤五实现。

[0325] 步骤一、电子设备基于该第一相似度和该第二相似度,确定第一损失值,该第一损失值与该第一相似度正相关,该第一损失值与该第二相似度负相关。

[0326] 该步骤一与上述点排序学习中的步骤一同理,在此不多做赘述。

[0327] 步骤二、电子设备根据该样本图像与该第二样本图像集合中样本图像之间的第三相似度,确定该第二样本图像集合中样本图像中该第一样本图像的第一中心,以及该第二样本图像的第二中心。

[0328] 在步骤二中,可以分别将正样本的分布中心作为第一中心,负样本的分布中心为第二中心。该第一中心和第二中心可以基于第三相似度确定。

[0329] 步骤三、电子设备根据该样本图像和该第二样本图像集合中样本图像分别与该第一中心以及该第二中心之间的距离,获取第二损失值。

[0330] 在确定出正样本的第一中心和负样本的第二中心之后,针对第二样本图像集合中样本图像,可以基于该第二样本图像集合中样本图像与两个中心之间的距离来确定出第二损失值。

[0331] 可以理解地,训练的初衷是使得正样本更加靠近第一中心,远离第二中心;负样本更加靠近第二中心,远离第一中心。对于一个锚点 z_i ,为了大幅度地分离其两个组距离分布,我们为其正负样本组设置两个较远的分布中心。然后,通过将每个分布压缩到各自设定的中心,可以使最小化分布的重叠。

[0332] 在一个具体示例中,对于正样本对和负样本对,电子设备可以基于下述公式八,确定第三相似度。

$$[0333] \quad d_{strc}(z_i, z_j) = \sqrt{\sum_{d=1}^D \omega_i^{(d)} \cdot (z_i^{(d)} - z_j^{(d)})^2}, \text{ 其中, } r_{ij} = z_i^{(d)} - z_j^{(d)}, \text{ 公式八其}$$

$$= \sqrt{\omega_i^T (r_{ij} \odot r_{ij})}$$

中, \odot 是哈达玛积。

[0334] 然后通过下述公式九确定第二损失值:

$$[0335] \quad \mathcal{L}_{strc} = \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i} c_{ij} |d_{strc}(z_i, z_j) - \eta_1| + (1 - c_{ij}) |d_{strc}(z_i, z_j) - \eta_2|, \text{ 公式九}$$

[0336] 其中,对于分布中心的软分配 c_{ij} ,可以通过下述公式十表示。

$$[0337] \quad c_{ij} = \frac{\exp(-\alpha |d_{strc}(z_i, z_j) - \eta_1|)}{\sum_{m=1}^2 \exp(-\alpha |d_{strc}(z_i, z_j) - \eta_m|)}, \text{ 公式十}$$

[0338] 其中, η_1 和 η_2 分别是第一中心和第二中心。

[0339] 例如,将距离分布可视化可以如图8所示,(a)中采用了未针对正负样本中心进行训练的图像之间距离分布情况,(b)中采用了正负样本中心进行训练的图像之间距离分布情况。通过训练,确保正样本和负样本之间的间隔较大且分布重叠较小,并改善了特征的鉴别性和泛化性。

[0340] 步骤四、电子设备基于该第一损失值和该第二损失值,对该第一特征提取网络进行训练。

[0341] 该步骤四与上述点排序学习中的步骤五同理,在此不多做赘述。

[0342] 通过组排序学习,将正样本、负样本一起进行训练,最终使得正样本和负样本之间的距离越来越大,正样本和负样本各自集中,这样也就说明第一特征提取网络和第二特征提取网络能够准确地提取到用于表示图像内容和邻域结构的第一特征和第二特征。

[0343] 在步骤S49中,电子设备基于该第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系,基于该第二预测相似度关系与目标相似度关系之间的差异,对该第二特征提取网络进行训练。

[0344] 除了第一特征提取网络,电子设备也可以对第二特征提取网络进行训练,在对第二特征提取网路进行训练时,可以使用第二特征得到的第三相似度进行训练即可。

[0345] 在一些实施例中,电子设备无需获取第一损失值,直接使用上述第二损失值,对该第二特征提取网络进行训练。第二损失值的获取过程同理,在此不多做赘述。

[0346] 该步骤S45至步骤S49为基于该样本图像的第一特征、第二特征以及该样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,基于该预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络和该第二特征提取网络进行训练的过程,不仅考虑到了样本图像本身的特征,还考虑到了该样本图像的邻域结构,从两个角度来对样本图像的图像特征进行处理,考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,基于训练后的特征提取网络进行图像检索也能够得到更加准确的检索结果。

[0347] 在一些实施例中,上述第一特征提取网络和第二特征提取网络训练完成后,还可以用于图像处理。具体地,图像处理过程可以为:电子设备响应于任一账号对图像的检索指令,基于训练得到的该第一特征提取网络对该图像进行处理,得到该图像的第一特征,基于该图像的第一特征确定该图像所匹配的目标图像,然后基于目标图像进行图像推送。

[0348] 在另一些实施例中,电子设备可以响应于任一账号对图像的检索指令,基于训练得到的该第一特征提取网络对该图像进行处理,得到该图像的第一特征,基于训练得到的该第二特征提取网络对该图像进行处理,得到该图像的第二特征,然后基于该图像的第一特征和第二特征,确定该图像所匹配的目标图像,然后基于目标图像进行图像推送。

[0349] 在一个具体的可能实施例中,电子设备响应于任一账号对图像的检索指令,获取该图像的图像特征,基于训练得到的该第一特征提取网络对该图像的图像特征进行处理,得到该图像的第一特征,基于该图像的第一特征与多个候选图像的第一特征之间的第三相似度,从该多个候选图像中,确定该第三相似度符合条件的多个第一目标图像,然后电子设备向该账号推送该多个第一目标图像。该第一特征提取网络训练时考虑到了样本图像的邻域结构的特征,因而网络参数更加准确,提取到的第一特征也就更加准确,基于此进行图像检索,能够有效提高检索结果的准确性。

[0350] 在一些实施例中,电子设备还可以再基于图像的邻域结构的特征进一步对图像检索结果进行重排序。具体地,电子设备可以基于训练得到的该第二特征提取网络对该图像的图像特征进行处理,得到该图像的第二特征,然后基于该图像的第二特征与每个该第一目标图像的第二特征之间的第四相似度,对该多个第一目标图像进行排序,得到该多个第一目标图像的排序结果,然后电子设备可以按照该排序结果向该账号推送该多个第一目标图像。

[0351] 本公开实施例,新增了一种用于提取第二特征的第二特征提取网络,第二特征用于表征目标类别图像的内容特征分布状况,也即是该样本图像的邻域结构,结合用于提取样本图像本身特征的第一特征提取网络,从两个角度来对样本图像的图像特征进行处理,

考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,这样提取到的特征能够更加准确地体现出图像所属的类别,基于训练后的特征提取网络进行图像检索也能够得到更加准确的检索结果。

[0352] 上述图5所示实施例,对特征提取网络的训练方法进行了重点说明,下面通过图9所示,对图像处理方法进行说明。图9是根据一示例性实施例示出的一种图像处理方法的流程图,如图9所示,该方法用于电子设备中,包括以下步骤。

[0353] 在步骤S81中,电子设备获取图像。

[0354] 在步骤S82中,电子设备对图像进行特征提取,得到该图像的图像特征。

[0355] 该步骤S81和步骤S82为获取图像的图像特征的过程,该过程与上述步骤S41和步骤S42同理,在此不多做赘述。

[0356] 在步骤S83中,电子设备对该图像的图像特征进行处理,得到该图像的第一特征,该第一特征用于表征该图像的内容特征。

[0357] 需要说明的是,该步骤S81至步骤S83为对该图像进行处理,得到该图像的第一特征的一种可能实现方式,在此仅以先获取图像,然后获取到图像的图像特征后,将其输入到第一特征提取网络中进行处理为例。在另一些实施例中,可以直接将图像输入第一特征提取网络中,由第一特征提取网络对其进行处理,得到第一特征。本公开实施例对具体采用哪种实现方式不作限定。

[0358] 在步骤S84中,电子设备对该图像的图像特征进行处理,得到该图像的第二特征,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该图像所属的类别。

[0359] 需要说明的是,该步骤S81、S82和步骤S84为对该图像进行处理,得到该图像的第二特征的一种可能实现方式,在此仅以先获取图像,然后获取到图像的图像特征后,将其输入到第二特征提取网络中进行处理为例。在另一些实施例中,可以直接将图像输入第二特征提取网络中,由第二特征提取网络对其进行处理,得到第二特征。本公开实施例对具体采用哪种实现方式不作限定。

[0360] 对于上述步骤S83和步骤S84可以分别通过第一特征提取网络和第二特征提取网络实现。也即是,电子设备可以基于第一特征提取网络对所述图像进行处理,得到所述第一特征,基于第二特征提取网络,对所述图像进行处理,得到所述第二特征。

[0361] 其中,所述第一特征提取网络和所述第二特征提取网络分别基于第一样本图像集合中样本图像之间的预测相似度关系与目标预测相似度关系之间的差异训练得到,所述预测相似度关系基于样本图像的第一特征、第二特征以及所述样本图像所属类别确定,且所述预测相似度关系用于表征两个样本图像之间是否属于同一类别。

[0362] 该第一特征提取网络和第二特征提取网络的训练过程具体可以参见上述图5所示实施例。在此对训练过程进行简单说明。

[0363] 在一些实施例中,所述第一特征提取网络和所述第二特征提取网络的训练过程可以通过下述步骤一至步骤七实现。

[0364] 步骤一、基于所述样本图像以及所述第一样本图像集合中第一样本图像的第一特征,确定所述样本图像与所述第一样本图像之间的第一相似度,所述第一样本图像为所述

样本图像集合中与所述样本图像属于相同类别的图像。

[0365] 步骤二、基于所述样本图像以及与所述第一样本图像集合中第二样本图像的第一特征,确定所述样本图像与所述第二样本图像之间的第二相似度,第二样本图像为所述第一样本图像集合中与所述样本图像属于不同类别的图像。

[0366] 步骤三、基于所述样本图像的第二特征、所述样本图像与所述第二样本图像集合中样本图像的第一特征,确定所述样本图像和所述第二样本图像集合中样本图像之间的第三相似度,所述第二样本图像集合包括所述第一样本图像和所述第二样本图像。

[0367] 步骤四、基于所述第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图像之间的第一预测相似度关系。

[0368] 步骤五、基于所述第三相似度,确定第一样本图像集合中样本图像之间的第二预测相似度关系。

[0369] 步骤六、基于所述第一预测相似度关系与目标相似度关系之间的差异,对所述第一特征提取网络进行训练。

[0370] 步骤七、基于所述第二预测相似度关系与目标相似度关系之间的差异,对所述第二特征提取网络进行训练。

[0371] 在一些实施例中,上述步骤三可以包括多种实现方式。

[0372] 实现方式一:根据所述第一相似度和所述第二相似度,从所述第二样本图像集合中,确定所述样本图像对应的第三样本图像集合,所述第三样本图像集合包括所述第一相似度或所述第二相似度满足条件的图像;以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征和所述第三样本图像集合中样本图像的第一特征之间的距离;对所述距离进行归一化,得到所述样本图像和所述第三样本图像集合中样本图像之间的第三相似度。

[0373] 在该实现方式一中,相应的,步骤四可以通过下述方式确定第一样本图像集合中样本图像之间的第一预测相似度关系:基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;基于所述样本图像与所述第三样本图像集合中样本图像之间的第三相似度,对所述第三样本图像集合中样本图像进行排序,得到预测排序结果。进而步骤六中在对第一特征提取网络进行训练时,电子设备可以基于所述样本图像与所述第三样本图像集合中样本图像之间的第一相似度,对所述第三样本图像集合中样本图像进行排序,得到目标排序结果;基于所述预测排序结果与目标排序结果,确定第二损失值,所述第二损失值用于指示所述排序结果和所述目标排序结果之间的差异;基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0374] 上述步骤三的实现方式二:以所述样本图像的第二特征作为系数,获取所述样本图像的第一特征与所述第二样本图像集合中样本图像的第一特征之间的距离,所述距离为所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度。

[0375] 相应的,上述步骤四和步骤六中,电子设备可以基于所述第一相似度和所述第二相似度,确定第一损失值,所述第一损失值与所述第一相似度正相关,所述第一损失值与所述第二相似度负相关;根据所述样本图像与所述第二样本图像集合中样本图像之间的第三相似度,确定所述第二样本图像集合中样本图像中所述第一样本图像的第一中心,以及所

述第二样本图像的第二中心;根据所述样本图像和所述第二样本图像集合中样本图像分别与所述第一中心以及所述第二中心之间的距离,获取第二损失值;基于所述第一损失值和所述第二损失值,对所述第一特征提取网络进行训练。

[0376] 该步骤S83和步骤S84可与上述步骤S43和步骤S44同理,对于第二特征,在一些实施例中,电子设备可以基于第二特征提取网络中的两个全连接层,对该图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该图像的第二特征。本公开实施例在此不做过多赘述。

[0377] 在步骤S85中,电子设备基于该图像的第一特征与候选图像集合中候选图像的第一特征之间的第四相似度,从该候选图像集合中,确定该图像所匹配的目标图像,该目标图像的该第四相似度符合条件。

[0378] 基于上述过程提取到图像的第一特征之后,则可以与候选图像进行匹配,匹配过程通过比较图像与候选图像之间的第四相似度来确定。该第四相似度的获取过程与上述第一相似度的获取过程同理,在此不作过多赘述。

[0379] 该第四相似度符合条件可以由相关技术人员根据需求进行设置,比如,该第四相似度符合条件为第四相似度大于相似度阈值,或者目标图像为第四相似度最大的目标数量的候选图像。

[0380] 在步骤S86中,电子设备基于该图像的第二特征与该目标图像的第二特征之间的第五相似度,对目标图像进行排序,得到该目标图像的排序结果。

[0381] 基于第一特征进行图像检索后,电子设备还可以考虑到图像的邻域结构特征,进行重新排序,以重新排序后的结果进行推荐。

[0382] 该步骤S85和步骤S86为基于该图像的该第一特征和该第二特征,确定该图像所匹配的目标图像的过程。在一些实施例中,上述目标图像匹配过程也可以无需确定第四相似度和第五相似度,而是直接基于第一特征和第二特征进行处理,得到第三特征,基于第三特征与候选图像集合中候选图像进行匹配,得到匹配的目标图像。

[0383] 在步骤S87中,电子设备按照该排序结果对该目标图像进行图像推送。

[0384] 需要说明的是,上述步骤S87为基于该目标图像进行图像推送的一种可能实现方式,上述仅以通过第一特征选出目标图像并基于第二特征对目标图像进行重排序的方式实现为例。在一些实施例中,上述目标图像匹配过程也可以无需确定第四相似度和第五相似度,而是直接基于第一特征和第二特征进行处理,得到第三特征,基于第三特征与候选图像集合中候选图像进行匹配,得到匹配的目标图像,然后基于目标图像进行推送。本公开实施例对具体采用哪种方式不作限定。

[0385] 上述图像处理过程中,使用了用于表示图像内容和邻域结构的特征,进行了二次图像检索,相较于仅通过图像内容进行检索的方式,还考虑到了图像的邻域结构,这样图像推送结果更加准确,更加符合用户需求。本公开实施例,新增了一种用于提取第二特征的第二特征提取网络,第二特征用于表示与样本图像属于同一类别的样本图像的分布,也即是该样本图像的邻域结构,结合用于提取样本图像本身特征的第一特征提取网络,从两个角度来对样本图像的图像特征进行处理,考虑到了不同样本图像的邻域结构可能不均匀的情况,在训练过程中不仅学习样本图像本身的特征,还能够自适应地学习到样本图像的邻域结构,以此训练得到的特征提取网络在特征提取方面的准确性更好,基于训练后的特征提

取网络进行图像检索也能够得到更加准确的检索结果。

[0386] 下面提供一个具体示例,该具体示例示出了本公开提供的方法应用于不同的数据集上的实验结果。

[0387] 具体地,针对二值样本嵌入,分别在CIFAR-10,NUS-WIDE和FLICKR25K三个大规模数据集上进行了实验。下面针对每个数据集以及实验设置进行介绍。

[0388] 数据集一: CIFAR-10

[0389] CIFAR-10是一个用于识别物体的十类别数据集。十个类别分别为:飞机、汽车、鸟类、猫、鹿、狗、蛙类、马、船和卡车。每个图片的尺寸为 32×32 ,每个类别有6000个图像,数据集中一共有50000张训练图片和10000张测试图片。采用5000张图像用于训练,1000张图像用于测试。

[0390] 数据集二: NUS-WIDE

[0391] NUS-WIDE是大规模的图像数据集,包含81个语义概念和269648幅网络图像。在实验时,选择了其中最大的21个语义概念,以及21个语义概念对应的186577幅图像进行实验。然后在其中随机选择了2100个样本图像作为测试集,其他作为检索集,并从检索集中抽取10500个样本图像作为训练集。

[0392] 数据集三: FLICKR25K

[0393] FLICKR25K是从Flickr中收集的25000个图像。类似于NUS-WIDE,在实验时,随机选择了2000个样本图像作为测试集,其他作为检索集,并从检索集中抽取5000个样本图像作为训练集。

[0394] 对于实值样本嵌入,采用的数据集分别是CUB-200-2011,Cars-196和Stanford Online Products。

[0395] 数据集一: CUB-200-2011拥有200种鸟类的11788张图像。使用前100种(5864张图像)进行训练,其余100种(5924张图像)进行测试。

[0396] 数据集二: Cars-196有196辆汽车的16185张图像。将最初的98辆汽车(8054张图像)进行了训练,将其余100辆汽车(8131张图像)进行了测试。

[0397] 数据集三: Stanford Online Products包含22634个在线产品的120053张图像。使用第一批11318产品(59551张图像)进行培训,并使用11316产品(60502张图像)进行测试。

[0398] 对于二值嵌入,我们在四个有代表性的二值度量方法上做了实验,分别是:SGH,SSDH,GreedyHash和TBH。对于实值嵌入,我们在三个有代表性的度量学习方法上做了实验,分别是:Triplet loss,Margin loss和MS loss。

[0399] 在测试阶段,我们直接使用样本嵌入(NASA)来检索粗粒度结果。然后,为进一步证明结构嵌入的有效性(NASA[†]),我们对召回的前 n 个结果进行重新排序,其中CIFAR-10为 $n=2000$,NUS-WIDE和FLICKR25K为 $n=10000$,CUB-200-2011和Cars-196为 $n=32$,Stanford Online Products为 $n=2000$ 。值得注意的是,由于 n 的值很小,我们的重新排序仅引入了很少的计算开销。

[0400] 最终实验结果如下表一、表二和表三所示,通过六个数据集实验表明,本公开提出的方法可以显著提升检索模型的效果。其中,本公开的方法在CUB-200-2011数据集上的表现可以如图10所示,相同类别的图像分布在一个邻域中。

[0401] 表一

[0402]

方法	骨干	CIFAR-10				NUS-WIDE				FLICKR25K			
		16bits (比特)	32bits	64bits	128bits	16bits	32bits	64bits	128bits	16bits	32bits	64bits	128bits
SGH	Alex	43.5	43.7	44.3	45.6	66.2	67.8	69.2	70.4	66.6	67.5	48.2	69.4
NASA+SGH	Net	43.9	45.8	46.8	48.1	67.2	68.7	70.5	71.1	68.6	69.3	70.0	70.7

[0403]

NASA [†] +SGH		44.7	46.1	47.3	49.8	68.1	69.8	71.9	72.2	69.2	70.1	70.5	71.5
SSDH	VGG	24.9	27.4	29.3	30.1	63.1	63.4	64.1	65.2	75.2	75.6	75.7	76.0
NASA+SSDH	-16	26.5	30.2	33.9	37.7	63.8	64.6	68.3	69.1	74.6	75.9	76.6	78.8
NASA [†] +SSDH		28.8	32.4	35.3	38.2	64.2	65.8	70.2	70.5	75.3	77.1	77.5	79.2
Greedy Hash	Alex	32.7	36.1	40.2	43.1	70.5	73.1	75.0	76.5	65.9	66.7	70.3	70.8
NASA+Greedy Hash	Net	34.9	39.4	43.6	46.0	73.3	76.5	78.4	79.6	71.0	73.0	74.2	74.8
NASA [†] +Greedy Hash		36.6	40.8	44.2	46.3	74.2	77.3	78.7	79.7	72.0	73.5	74.3	74.8
TBH	Alex	48	51.4	53.2	54.2	71.7	72.5	73.5	73.7	69.5	70.1	71.5	71.1
NASA+TBH	Net	48	51.9	53.7	55.5	72.4	72.9	75.8	76.1	69.9	70.8	72.1	72.5
NASA [†] +TBH		48.1	52.2	54.3	55.8	73.8	74.0	76.3	76.6	70.3	71.5	73.1	73.6

[0404] 如表一所示,表一中的数值均为MAP (Mean Average Precision,平均检索精度),MAP越高则说明越准确。

[0405] 表二

[0406]

方法	CUB-200-2011				Cars-196				Stanford Online Products			
	R@1	R@2	R@4	R@8	R@1	R@2	R@4	R@8	R@1	R@10	R@10 ²	R@10 ³
Triplet ₁₂₈	55.5	67.1	77.5	85.4	75.8	84.3	90.1	93.4	72.4	86.5	95.0	98.1
NASA _{128+Triplet}	64.7	76.1	84.4	90.2	85.3	91.0	94.4	96.7	77.1	89.5	95.4	98.5
NASA [†] _{128+Triplet}	65.5	76.5	84.6	90.3	85.8	91.4	94.8	96.9	77.4	89.8	95.7	98.6
Margin ₁₂₈	62.8	73.9	82.4	89.7	78.1	86.2	91.1	94.2	72.7	86.2	93.8	98.0

[0407]

NASA _{128+Margin}	65.1	76.3	84.3	90.5	82.6	89.1	93.0	95.6	74.2	87.4	94.5	98.2
NASA [†] _{128+Margin}	65.6	76.5	84.4	90.9	83.1	89.2	93.1	95.7	74.4	87.8	94.6	98.2
MS ₁₂₈	61.7	72.8	82.7	90.2	81.1	89.0	93.4	96.4	76.5	89.3	95.1	98.4
NASA _{128+MS}	64.3	75.4	84.4	90.7	85.0	90.8	94.5	96.9	77.2	89.5	95.6	98.7
NASA [†] _{128+MS}	64.8	75.6	84.7	90.8	85.3	90.9	94.7	97.1	77.3	89.8	95.7	98.7

[0408] 如表二所示,表二中的数值均为召回率,本申请提供的方法的召回率,明显要高于其他方法的召回率。

[0409] 表三

[0410]

方法	骨干	参考	CUB-200-2011				Cars-196				Stanford Online Products			
			R@1	R@2	R@4	R@8	R@1	R@2	R@4	R@8	R@1	R@10	R@10 ²	R@10 ³
Smart Mining ⁶⁴	G	ICC V17	49.8	62.3	74.1	83.3	64.7	76.2	84.2	90.2	-	-	-	-
Proxy NCA ¹²⁸	B N	ICC V17	58.3	70.9	78.1	86.8	78.3	85.9	90.2	93.1	-	-	-	-
A-BIE R ⁵¹²	G	ICC V18	57.5	68.7	78.3	86.2	82.0	89.0	93.2	96.1	74.2	86.9	94.0	97.8
MIC ⁵¹²	R	ICC V19	66.1	76.8	85.6	-	82.6	89.1	93.2	-	77.2	89.4	95.6	-
SoftTriple ⁵¹²	B N	ICC V19	65.4	76.4	84.5	90.4	84.5	90.7	94.5	96.9	78.3	90.3	95.9	-
EE-MS ⁵¹²	G	ICC V20	57.4	68.7	79.5	86.9	75.8	83.9	89.7	93.8	78.1	90.3	95.8	-
XBM ⁵¹	G	ICC	61.9	72.9	81.2	88.6	80.3	87.1	91.9	95.1	77.4	89.6	95.4	98.4

[0411]

		V20												
Circle ⁵¹²	B N	ICC V20	66.7	77.4	86.2	91.2	83.4	89.8	94.1	96.5	78.3	90.5	96.1	98.6
Proxy Anchor ⁵¹²	B N	ICC V20	68.4	79.2	86.8	91.6	86.1	91.7	95.0	97.3	79.1	90.8	96.2	98.7
NASA ⁵¹² +Triple	B N	Proposed	68.3	79.1	86.6	91.5	88.3	93.1	95.9	97.6	78.1	90.3	96.0	98.7
NASA ⁵¹² +Triple	B N	Proposed	68.8	79.3	86.9	91.9	89.0	93.4	96.2	97.8	78.5	90.6	96.2	98.8

[0412] 表三中的数值均为召回率,本申请提供的方法的召回率,明显要高于其他方法的召回率。

[0413] 需要说明的是,本公开提供的图像处理模型的确定方法能够用于任意图像检索场景。例如,可以以软件的方式安装于用户的工作电脑,提供实时哈希检索,也可以将其安装于后台服务器,提供大批量后台检索。该图像检索场景可以为商品图像检索、医学图像检索等等,当然,也可以为对包含有其他搜索目标的图像的检索场景,例如,在商品图像检索场景中,可以用于日常生活中对未知商品的快速检索。这样无需寻找专家,用户可以为商品拍照,通过本公开提供的方法,就可以立即检索到商品的精准信息。又例如,在一种图像检索场景中,比如城市管理场景中,能够通过上述方式对采集到包含有任一目标的图像进行精确分析,快速搜索到搜索目标。借助城市大规模信息采集系统,让计算机根据采集数据快速搜索到搜索目标在何时何地出现。该搜索目标可以为满足搜索条件的人员,也可以为其他生物或物品,比如,进行过某种行为的人员、走失人员、走失宠物、符合条件的车辆等,从而为保障城市管理和安全提供精准的数据参考。又例如,在医学图像检索场景中,能够用于医学诊断辅助,医院每天都会产生大量的医学图像数据。高效的组织、管理和检索大规模的医学图像数据,能够辅助医师对不同疾病进行精确的诊断,有着重要的临床价值和商业前景。通过上述方式能够对医学图像进行分析,以确定出医学图像中的人体组织分布等,为疾病诊断提供辅助。

[0414] 图11是根据一示例性实施例示出的一种特征提取网络训练装置框图。参照图11,该装置包括:

[0415] 处理单元1001,被配置为执行基于第一特征提取网络对样本图像进行处理,得到该样本图像的第一特征,该第一特征用于表征该样本图像的内容特征;

[0416] 该处理单元1001,还被配置为执行基于第二特征提取网络对该样本图像进行处理,得到该样本图像的第二特征,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该样本图像所属的类别;

[0417] 确定单元1002,被配置为执行基于该样本图像的第一特征、第二特征以及该样本图像所属类别,确定第一样本图像集合中样本图像之间的预测相似度关系,该第一样本图像集合包括属于同一类别的样本图像和属于不同类别的样本图像,该预测相似度关系用于表征两个样本图像之间是否属于同一类别;

[0418] 训练单元1003,被配置为执行基于该预测相似度关系与目标相似度关系之间的差异,对该第一特征提取网络和该第二特征提取网络进行训练。

[0419] 在一些实施例中,该确定单元1002被配置为执行:

[0420] 基于该样本图像以及该第一样本图像集合中第一样本图像的第一特征,确定该样本图像与该第一样本图像之间的第一相似度,该第一样本图像为该样本图像集合中与该样本图像属于相同类别的图像;

[0421] 基于该样本图像以及与该第一样本图像集合中第二样本图像的第一特征,确定该样本图像与该第二样本图像之间的第二相似度,第二样本图像为该第一样本图像集合中与该样本图像属于不同类别的图像;

[0422] 基于该样本图像的第二特征、该样本图像与该第二样本图像集合中样本图像的第一特征,确定该样本图像和该第二样本图像集合中样本图像之间的第三相似度,该第二样本图像集合包括该第一样本图像和该第二样本图像;

[0423] 基于该第一相似度、第二相似度和第三相似度,确定第一样本图像集合中样本图

像之间的第一预测相似度关系；

[0424] 基于该第三相似度，确定第一样本图像集合中样本图像之间的第二预测相似度关系；

[0425] 该训练单元1003被配置为执行：

[0426] 基于该第一预测相似度关系与目标相似度关系之间的差异，对该第一特征提取网络进行训练；

[0427] 基于该第二预测相似度关系与目标相似度关系之间的差异，对该第二特征提取网络进行训练。

[0428] 在一些实施例中，该确定单元1002被配置为执行：

[0429] 根据该第一相似度和该第二相似度，从该第二样本图像集合中，确定该样本图像对应的第三样本图像集合，该第三样本图像集合包括该第一相似度或该第二相似度满足条件的图像；

[0430] 以该样本图像的第二特征作为系数，获取该样本图像的第一特征和该第三样本图像集合中样本图像的第一特征之间的距离；

[0431] 对该距离进行归一化，得到该样本图像和该第三样本图像集合中样本图像之间的第三相似度。

[0432] 在一些实施例中，该确定单元1002被配置为执行：

[0433] 基于该第一相似度和该第二相似度，确定第一损失值，该第一损失值与该第一相似度正相关，该第一损失值与该第二相似度负相关；

[0434] 基于该样本图像与该第三样本图像集合中样本图像之间的第三相似度，对该第三样本图像集合中样本图像进行排序，得到预测排序结果；

[0435] 该训练单元1003被配置为执行：基于该样本图像与该第三样本图像集合中样本图像之间的第一相似度，对该第三样本图像集合中样本图像进行排序，得到目标排序结果；

[0436] 基于该预测排序结果与目标排序结果，确定第二损失值，该第二损失值用于指示该排序结果和该目标排序结果之间的差异；

[0437] 基于该第一损失值和该第二损失值，对该第一特征提取网络进行训练。

[0438] 在一些实施例中，该确定单元1002被配置为执行：

[0439] 以该样本图像的第二特征作为系数，获取该样本图像的第一特征与该第二样本图像集合中样本图像的第一特征之间的距离，该距离为该样本图像与该第二样本图像集合中样本图像之间的第三相似度。

[0440] 在一些实施例中，该确定单元1002和该训练单元1003被配置为执行：

[0441] 基于该第一相似度和该第二相似度，确定第一损失值，该第一损失值与该第一相似度正相关，该第一损失值与该第二相似度负相关；

[0442] 根据该样本图像与该第二样本图像集合中样本图像之间的第三相似度，确定该第二样本图像集合中样本图像中该第一样本图像的第一中心，以及该第二样本图像的第二中心；

[0443] 根据该样本图像和该第二样本图像集合中样本图像分别与该第一中心以及该第二中心之间的距离，获取第二损失值；

[0444] 基于该第一损失值和该第二损失值，对该第一特征提取网络进行训练。

- [0445] 在一些实施例中,该第二特征提取网络包括两个全连接层和激活层;
- [0446] 该处理单元1001被配置为执行基于该第二特征提取网络中的该两个全连接层,对该样本图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该样本图像的第二特征。
- [0447] 图12是根据一示例性实施例示出的一种图像处理装置框图。参照图12,该装置包括:
- [0448] 处理单元1101,被配置为执行对该图像进行处理,得到该图像的第一特征,该第一特征用于表征该图像的内容特征;
- [0449] 该处理单元1101,还被配置为执行对该图像进行处理,得到该图像的第二特征,该第二特征用于表征目标类别图像的内容特征分布状况,该目标类别为该图像所属的类别;
- [0450] 确定单元1102,被配置为执行基于该图像的该第一特征和该第二特征,确定该图像所匹配的目标图像;
- [0451] 推送单元1103,被配置为执行基于该目标图像进行图像推送。
- [0452] 在一些实施例中,该确定单元1102被配置为执行:
- [0453] 基于该图像的第一特征与候选图像集合中候选图像的第一特征之间的第四相似度,从该候选图像集合中,确定该图像所匹配的目标图像,该目标图像的该第四相似度符合条件;
- [0454] 该推送单元1103被配置为执行:
- [0455] 基于该图像的第二特征与该目标图像的第二特征之间的第五相似度,对该目标图像进行排序,得到该目标图像的排序结果;
- [0456] 按照该排序结果对该目标图像进行图像推送。
- [0457] 在一些实施例中,所述处理单元1101被配置为执行:
- [0458] 基于第一特征提取网络对所述图像进行处理,得到所述第一特征;
- [0459] 所述处理单元1101被配置为执行:
- [0460] 基于第二特征提取网络,对所述图像进行处理,得到所述第二特征;
- [0461] 其中,所述第一特征提取网络和所述第二特征提取网络分别基于第一样本图像集合中样本图像之间的预测相似度关系与目标预测相似度关系之间的差异训练得到,所述预测相似度关系基于样本图像的第一特征、第二特征以及所述样本图像所属类别确定,且所述预测相似度关系用于表征两个样本图像之间是否属于同一类别。
- [0462] 在一些实施例中,该处理单元1101被配置为执行:
- [0463] 基于第二特征提取网络中的两个全连接层,对该图像进行前向计算,由该第二特征提取网络的激活层对最后一个全连接层的输出进行处理,输出该图像的第二特征。
- [0464] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。
- [0465] 图13是根据一示例性实施例示出根据一示例性实施例示出的一种电子设备的结构示意图,该电子设备1200可因配置或性能不同而产生比较大的差异,包括一个或一个以上处理器(Central Processing Units,CPU)1201和一个或一个以上的存储器1202,其中,该存储器1202中存储有至少一条指令,该至少一条指令由该处理器1201加载并执行以实现上述各个方法实施例提供的特征提取网络训练方法或图像处理方法。该电子设备还包括其

他用于实现设备功能的部件,例如,该电子设备还具有有线或无线网络接口以及输入输出接口等部件,以便进行输入输出。本申请实施例在此不做赘述。

[0466] 上述方法实施例中的电子设备实现为终端。例如,图14是根据一示例性实施例示出的一种终端的结构框图。该终端1300可以是便携式移动终端,比如:智能手机、平板电脑、MP3 (Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3) 播放器、MP4 (Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4) 播放器、笔记本电脑或台式电脑。终端1300还可能被称为用户设备、便携式终端、膝上型终端、台式终端等其他名称。

[0467] 通常,终端1300包括有:处理器1301和存储器1302。

[0468] 处理器1301可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处理器1301可以采用DSP (Digital Signal Processing,数字信号处理)、FPGA (Field-Programmable Gate Array,现场可编程门阵列)、PLA (Programmable Logic Array,可编程逻辑阵列) 中的至少一种硬件形式来实现。处理器1301也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理单元,也称CPU (Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器1301可以集成有GPU (Graphics Processing Unit,图像处理器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器1301还可以包括AI (Artificial Intelligence,人工智能) 处理器,该AI处理器用于处理有关机器学习的计算操作。

[0469] 存储器1302可以包括一个或多个计算机可读存储介质,该计算机可读存储介质可以是非暂态的。存储器1302还可包括高速随机存取存储器,以及非易失性存储器,比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中,存储器1302中的非暂态的计算机可读存储介质用于存储至少一个指令,该至少一个指令用于被处理器1301所执行以实现本申请中方法实施例提供的特征提取网络训练方法或图像处理方法。

[0470] 在一些实施例中,终端1300还可选包括有:外围设备接口1303和至少一个外围设备。处理器1301、存储器1302和外围设备接口1303之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口1303相连。具体地,外围设备包括:射频电路1304、显示屏1305、摄像头组件1306、音频电路1307、定位组件1308和电源1309中的至少一种。

[0471] 外围设备接口1303可被用于将I/O (Input/Output,输入/输出) 相关的至少一个外围设备连接到处理器1301和存储器1302。在一些实施例中,处理器1301、存储器1302和外围设备接口1303被集成在同一芯片或电路板上;在一些其他实施例中,处理器1301、存储器1302和外围设备接口1303中的任意一个或两个可以在单独的芯片或电路板上实现,本实施例对此不加以限定。

[0472] 射频电路1304用于接收和发射RF (Radio Frequency,射频) 信号,也称电磁信号。射频电路1304通过电磁信号与通信网络以及其他通信设备进行通信。射频电路1304将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频电路1304包括:天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路1304可以通过至少一种无线通信协议来与其它

终端进行通信。该无线通信协议包括但不限于：万维网、城域网、内联网、各代移动通信网络(2G、3G、4G及5G)、无线局域网和/或WiFi (Wireless Fidelity, 无线保真) 网络。在一些实施例中, 射频电路1304还可以包括NFC (Near Field Communication, 近距离无线通信) 有关的电路, 本申请对此不加以限定。

[0473] 显示屏1305用于显示UI (User Interface, 用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏1305是触摸显示屏时, 显示屏1305还具有采集在显示屏1305的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器1301进行处理。此时, 显示屏1305还可以用于提供虚拟按钮和/或虚拟键盘, 也称软按钮和/或软键盘。在一些实施例中, 显示屏1305可以为一个, 设置在终端1300的前面板; 在另一些实施例中, 显示屏1305可以为至少两个, 分别设置在终端1300的不同表面或呈折叠设计; 在另一些实施例中, 显示屏1305可以是柔性显示屏, 设置在终端1300的弯曲表面上或折叠面上。甚至, 显示屏1305还可以设置成非矩形的不规则图形, 也即异形屏。显示屏1305可以采用LCD (Liquid Crystal Display, 液晶显示屏)、OLED (Organic Light-Emitting Diode, 有机发光二极管) 等材质制备。

[0474] 摄像头组件1306用于采集图像或视频。可选地, 摄像头组件1306包括前置摄像头和后置摄像头。通常, 前置摄像头设置在终端的前面板, 后置摄像头设置在终端的背面。在一些实施例中, 后置摄像头为至少两个, 分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种, 以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR (Virtual Reality, 虚拟现实) 拍摄功能或者其它融合拍摄功能。在一些实施例中, 摄像头组件1306还可以包括闪光灯。闪光灯可以是单色温闪光灯, 也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合, 可以用于不同色温下的光线补偿。

[0475] 音频电路1307可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波, 并将声波转换为电信号输入至处理器1301进行处理, 或者输入至射频电路1304以实现语音通信。出于立体声采集或降噪的目的, 麦克风可以为多个, 分别设置在终端1300的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器1301或射频电路1304的电信号转换为声波。扬声器可以是传统的薄膜扬声器, 也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时, 不仅可以将电信号转换为人类可听见的声波, 也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中, 音频电路1307还可以包括耳机插孔。

[0476] 定位组件1308用于定位终端1300的当前地理位置, 以实现导航或LBS (Location Based Service, 基于位置的服务)。定位组件1308可以是基于美国的GPS (Global Positioning System, 全球定位系统)、中国的北斗系统或俄罗斯的伽利略系统的定位组件。

[0477] 电源1309用于为终端1300中的各个组件进行供电。电源1309可以是交流电、直流电、一次性电池或可充电电池。当电源1309包括可充电电池时, 该可充电电池可以是有线充电电池或无线充电电池。有线充电电池是通过有线线路充电的电池, 无线充电电池是通过无线线圈充电的电池。该可充电电池还可以用于支持快充技术。

[0478] 在一些实施例中, 终端1300还包括有一个或多个传感器1310。该一个或多个传感

器1310包括但不限于：加速度传感器1311、陀螺仪传感器1312、压力传感器1313、指纹传感器1314、光学传感器1315以及接近传感器1316。

[0479] 加速度传感器1311可以检测以终端1300建立的坐标系的三个坐标轴上的加速度大小。比如，加速度传感器1311可以用于检测重力加速度在三个坐标轴上的分量。处理器1301可以根据加速度传感器1311采集的重力加速度信号，控制显示屏1305以横向视图或纵向视图进行用户界面的显示。加速度传感器1311还可以用于游戏或者用户的运动数据的采集。

[0480] 陀螺仪传感器1312可以检测终端1300的机体方向及转动角度，陀螺仪传感器1312可以与加速度传感器1311协同采集用户对终端1300的3D动作。处理器1301根据陀螺仪传感器1312采集的数据，可以实现如下功能：动作感应（比如根据用户的倾斜操作来改变UI）、拍摄时的图像稳定、游戏控制以及惯性导航。

[0481] 压力传感器1313可以设置在终端1300的侧边框和/或显示屏1305的下层。当压力传感器1313设置在终端1300的侧边框时，可以检测用户对终端1300的握持信号，由处理器1301根据压力传感器1313采集的握持信号进行左右手识别或快捷操作。当压力传感器1313设置在显示屏1305的下层时，由处理器1301根据用户对显示屏1305的压力操作，实现对UI界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

[0482] 指纹传感器1314用于采集用户的指纹，由处理器1301根据指纹传感器1314采集到的指纹识别用户的身份，或者，由指纹传感器1314根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时，由处理器1301授权该用户执行相关的敏感操作，该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器1314可以被设置在终端1300的正面、背面或侧面。当终端1300上设置有物理按键或厂商Logo时，指纹传感器1314可以与物理按键或厂商Logo集成在一起。

[0483] 光学传感器1315用于采集环境光强度。在一个实施例中，处理器1301可以根据光学传感器1315采集的环境光强度，控制显示屏1305的显示亮度。具体地，当环境光强度较高时，调高显示屏1305的显示亮度；当环境光强度较低时，调低显示屏1305的显示亮度。在另一个实施例中，处理器1301还可以根据光学传感器1315采集的环境光强度，动态调整摄像头组件1306的拍摄参数。

[0484] 接近传感器1316，也称距离传感器，通常设置在终端1300的前面板。接近传感器1316用于采集用户与终端1300的正面之间的距离。在一个实施例中，当接近传感器1316检测到用户与终端1300的正面之间的距离逐渐变小时，由处理器1301控制显示屏1305从亮屏状态切换为息屏状态；当接近传感器1316检测到用户与终端1300的正面之间的距离逐渐变大时，由处理器1301控制显示屏1305从息屏状态切换为亮屏状态。

[0485] 本领域技术人员可以理解，图14中示出的结构并不构成对终端1300的限定，可以包括比图示更多或更少的组件，或者组合某些组件，或者采用不同的组件布置。

[0486] 上述方法实施例中的电子设备实现为服务器。例如，图15是根据一示例性实施例示出的一种服务器的结构示意图，该服务器1400可因配置或性能不同而产生比较大的差异，包括一个或一个以上处理器（Central Processing Units, CPU）1401和一个或一个以上的存储器1402，其中，所述存储器1402中存储有至少一条指令，所述至少一条指令由所述处

处理器1401加载并执行以实现上述各个方法实施例提供的特征提取网络训练方法或图像处理方法。当然,该服务器还具有有线或无线网络接口以及输入输出接口等部件,以便进行输入输出,该服务器还包括其他用于实现设备功能的部件,在此不做赘述。

[0487] 在示例性实施例中,还提供了一种计算机可读存储介质,例如包括至少一条指令的存储器,上述至少一条指令由可由电子设备的处理器执行以完成上述实施例中的特征提取网络训练方法或图像处理方法。可选地,计算机可读存储介质是只读存储器(Read-Only Memory,简称:ROM)、随机存取存储器(Random Access Memory,简称:RAM)、只读光盘(Compact Disc Read-Only Memory,简称:CD-ROM)、磁带、软盘和光数据存储设备等。

[0488] 在示例性实施例中,还提供了一种计算机程序产品,所述计算机程序产品或所述计算机程序包括一条或多条指令,所述一条或多条指令存储在计算机可读存储介质中。电子设备的一个或多个处理器从计算机可读存储介质中读取所述一条或多条指令,所述一个或多个处理器执行所述一条或多条指令,使得电子设备执行上述特征提取网络训练方法或图像处理方法。

[0489] 在一些实施例中,本申请实施例所涉及的计算机程序可被部署在一个电子设备上执行,或者在位于一个地点的多个电子设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个电子设备上执行,分布在多个地点且通过通信网络互连的多个电子设备可以组成区块链系统。

[0490] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其它实施方案。本申请旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0491] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。



图1

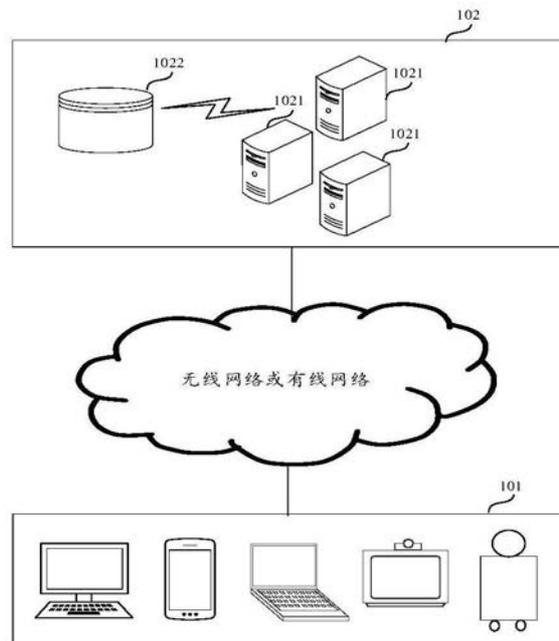


图2

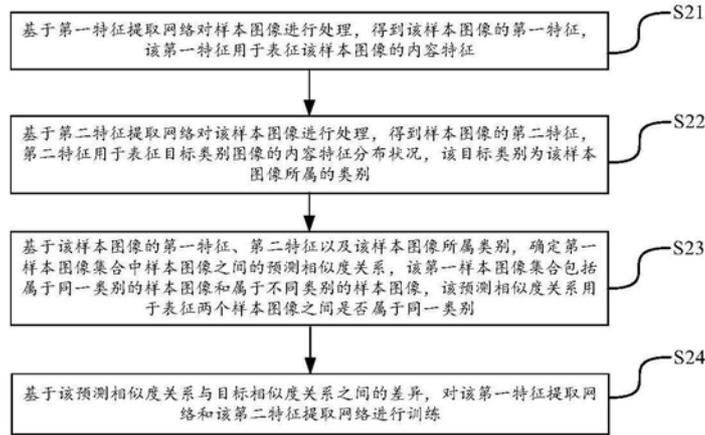


图3

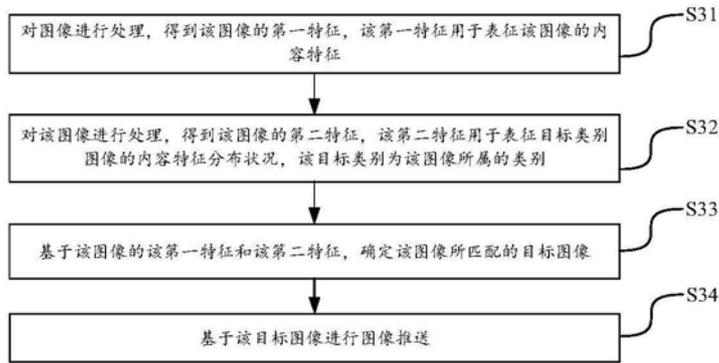


图4

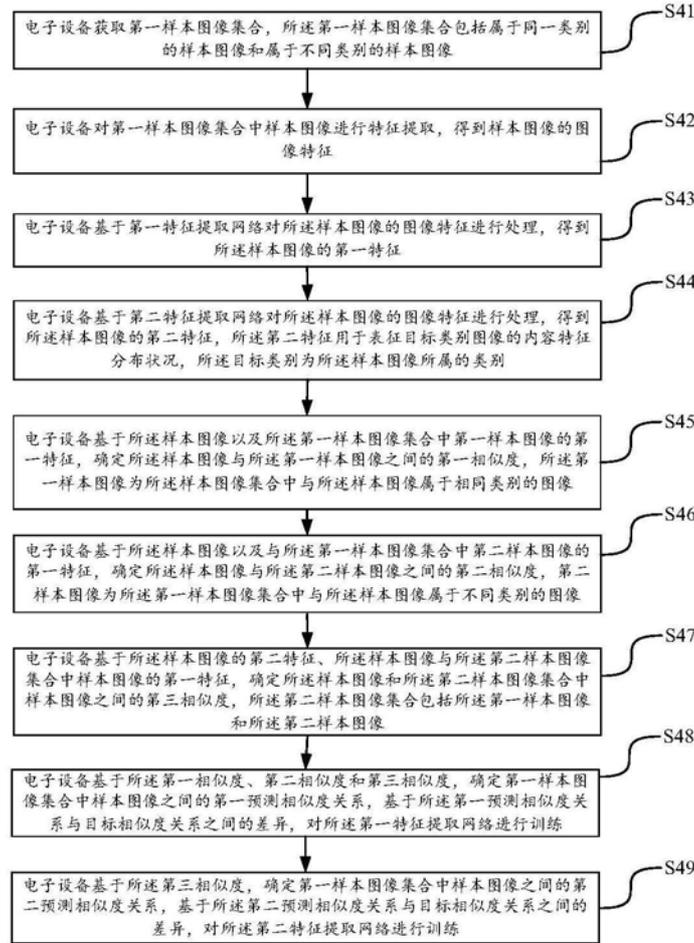


图5

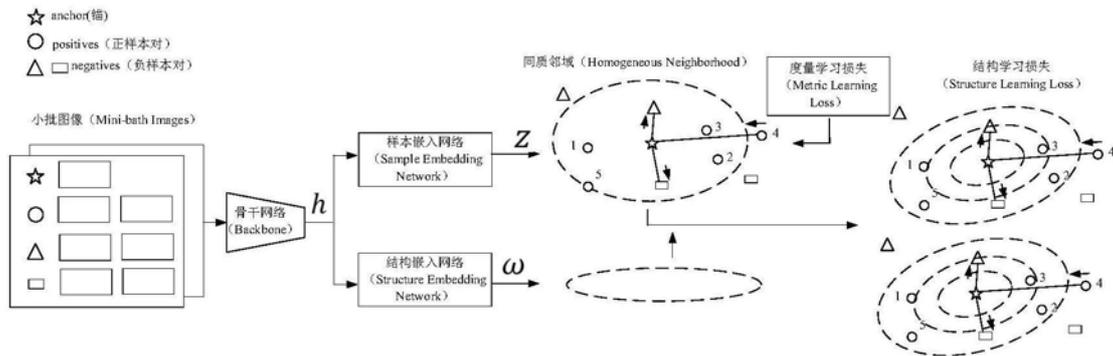


图6

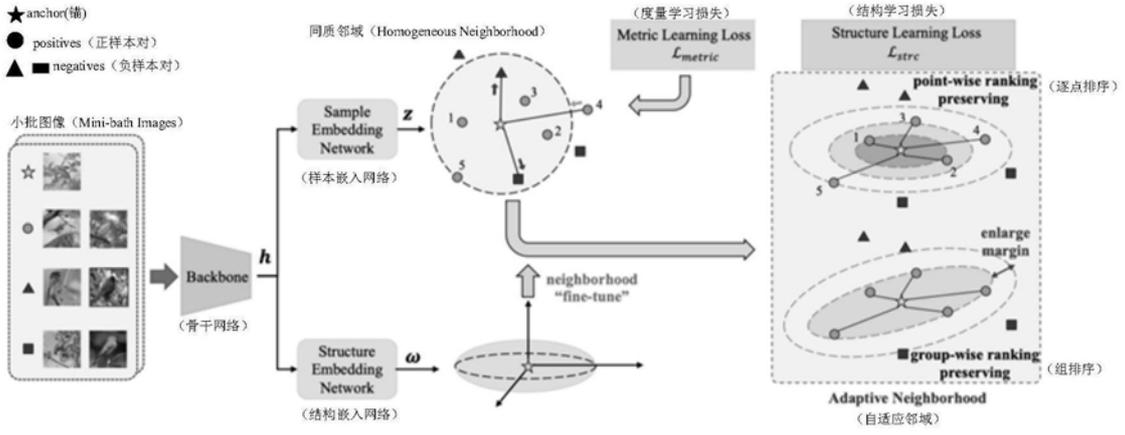


图7

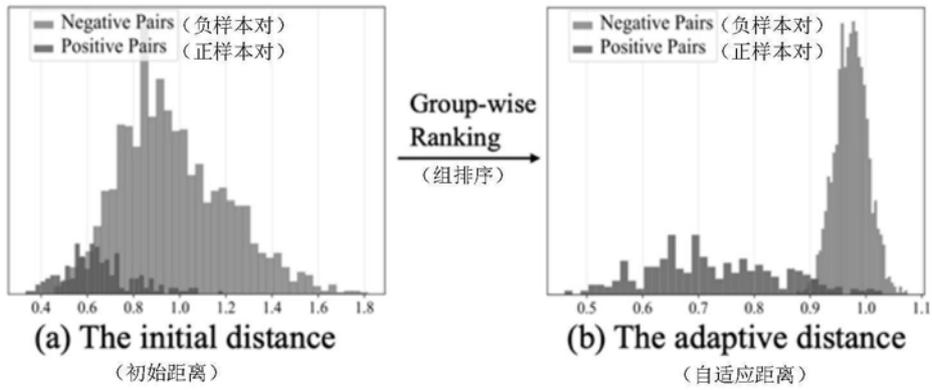


图8

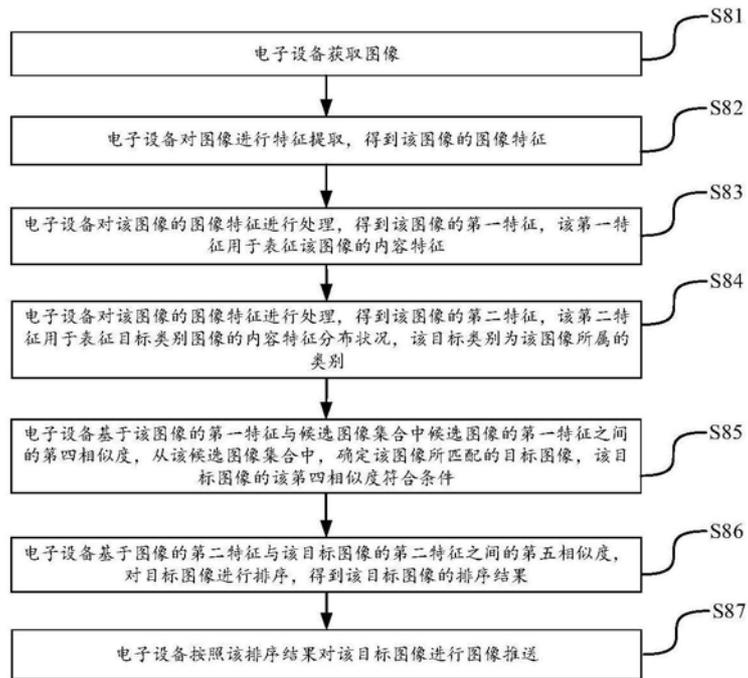


图9

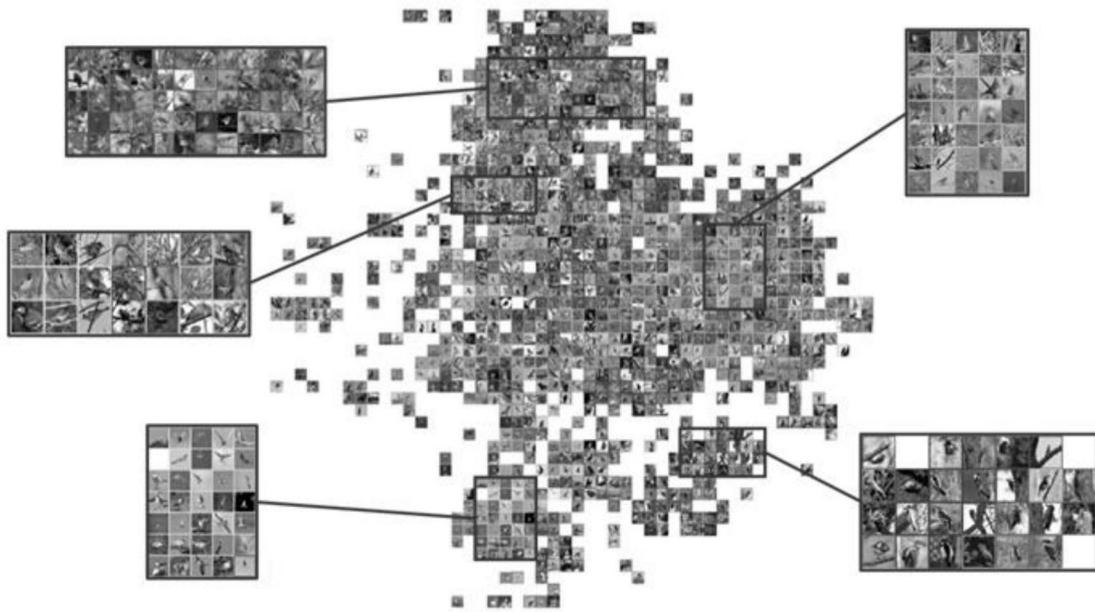


图10

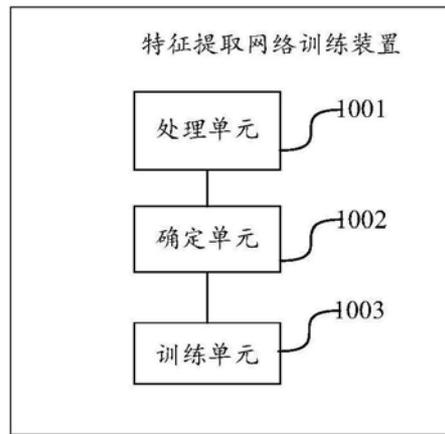


图11

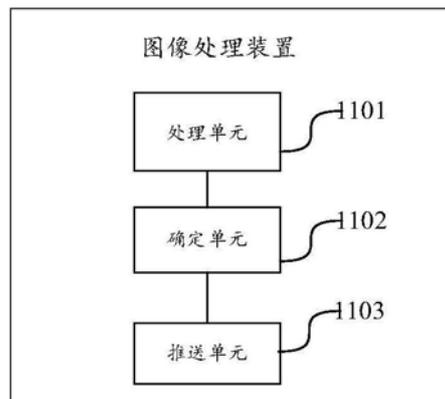


图12

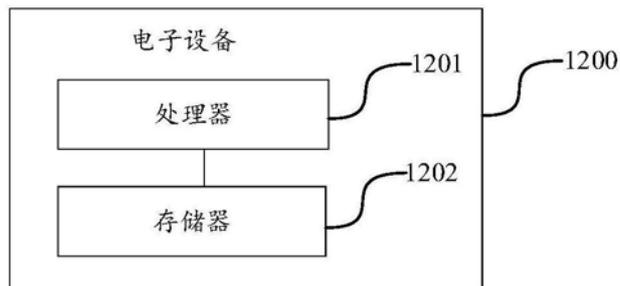


图13

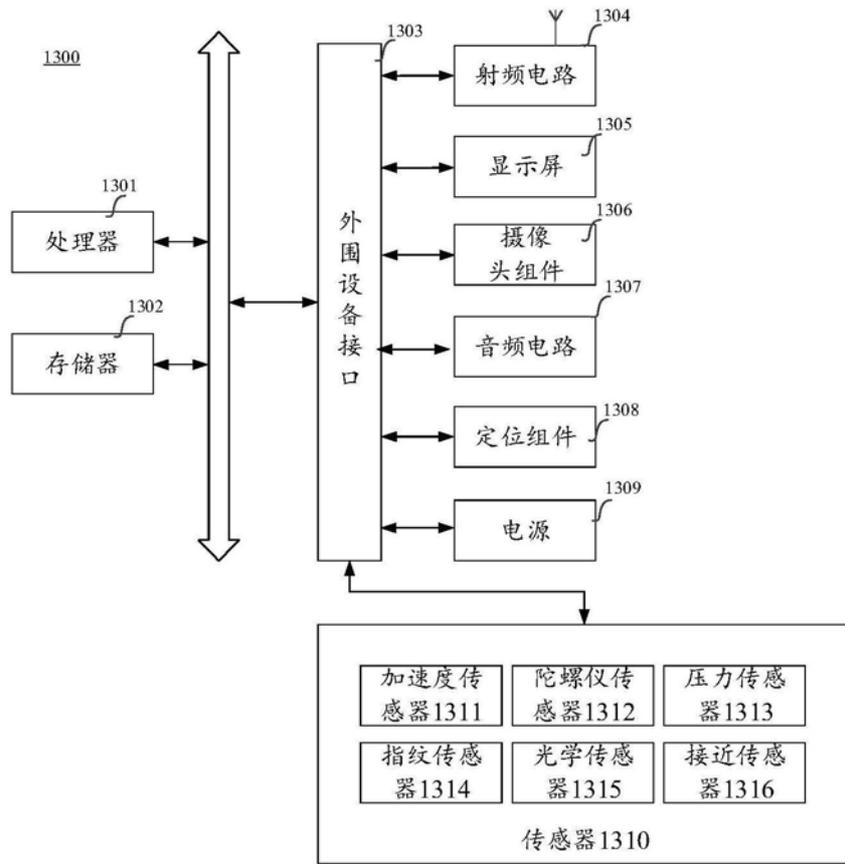


图14

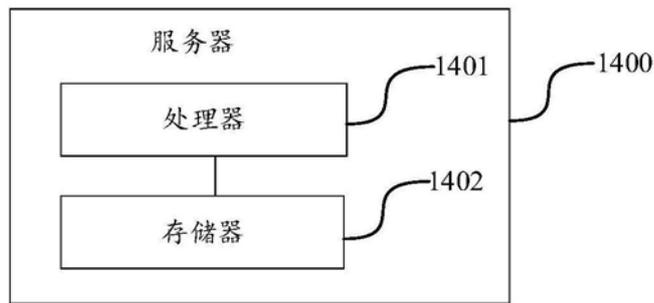


图15