



## (12) 发明专利申请

(10) 申请公布号 CN 112990348 A

(43) 申请公布日 2021.06.18

(21) 申请号 202110389258.5

(22) 申请日 2021.04.12

(71) 申请人 华南理工大学

地址 510630 广东省广州市天河区五山路  
381号

(72) 发明人 黄韬 余卫宇

(74) 专利代理机构 北京睿智保诚专利代理事务  
所(普通合伙) 11732

代理人 韩迎之

(51) Int.Cl.

G06K 9/62 (2006.01)

G06K 9/32 (2006.01)

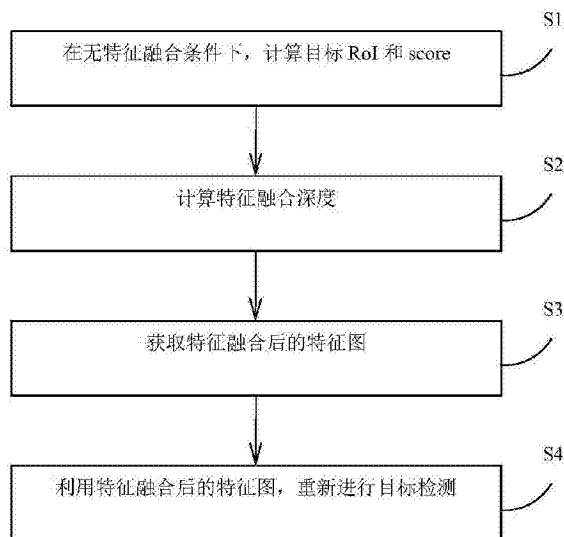
权利要求书1页 说明书4页 附图1页

### (54) 发明名称

一种自调节特征融合的小目标检测方法

### (57) 摘要

本发明公开了一种自调节特征融合的小目标检测方法,涉及人工智能领域。本发明包括以下步骤:获取第一特征图,利用目标检测算法,计算第一目标回归结果和第一分类结果,并通过第一目标回归结果和第一分类结果,得到感兴趣区域信息和第一特征图的置信度;感兴趣区域信息为感兴趣区域的高度和宽度;利用感兴趣区域的宽度和高度计算特征融合深度最大值;利用置信度计算特征融合深度;根据特征融合深度最大值和特征融合深度获取特征融合后的第二特征图;将第二特征图利用目标检测算法,计算第二目标回归和第二分类结果,完成目标检测。本发明综合考虑先验和后验信息以提高特征获取的准确度,减少了对“大小较小但容易检测的目标”的冗余消耗。



1. 一种自调节特征融合的小目标检测方法, 其特征在于, 包括以下步骤: 获取第一特征图, 利用目标检测算法, 计算第一目标回归结果和第一分类结果, 并通过所述第一目标回归结果和所述第一分类结果, 得到感兴趣区域信息和所述第一特征图的置信度score; 所述感兴趣区域信息为感兴趣区域的高度和宽度;

利用所述感兴趣区域的宽度和高度计算特征融合深度最大值; 利用所述置信度score计算特征融合深度;

根据所述特征融合深度最大值和所述特征融合深度利用特征金字塔获取特征融合后的第二特征图;

将第二特征图利用所述目标检测算法, 计算第二目标回归和第二分类结果, 完成目标检测。

2. 根据权利要求1所述的一种自调节特征融合的小目标检测方法, 其特征在于, 计算特征融合深度最大值具体如下:

$$d_{max} = \lfloor -\log_2 (\sqrt{h*w}/N_i) \rfloor ;$$

$N_i$  表示标准图像输入边长,  $h$  是感兴趣区域的高度,  $w$  是感兴趣区域的宽度, 且  $h, w \leq N_i$ 。

3. 根据权利要求1所述的一种自调节特征融合的小目标检测方法, 其特征在于, 计算特征融合深度具体如下:

$$d = \text{sum} (D(S_1, S_i) \leq \theta \cdot \lambda) ;$$

其中  $S_i$  表示分类置信度score按从大到小顺序排列中第  $i$  位的值,  $D(\cdot)$  表示  $S_1$  与  $S_i$  的距离, 在此取  $D(S_1, S_i) = |S_1 - S_i|$ ,  $\lambda$  表示阈值常量。

4. 根据权利要求1所述的一种自调节特征融合的小目标检测方法, 其特征在于, 特征金字塔获取特征图的具体过程如下:

第一特征图输入到特征金字塔中, 完成特征提取, 各特征提取层输出不同的特征, 经特征融合后输出多种特征图, 根据所述特征融合深度最大值和所述特征融合深度选取第二特征图。

## 一种自调节特征融合的小目标检测方法

### 技术领域

[0001] 本发明涉及人工智能领域,更具体的说是涉及一种自调节特征融合的小目标检测方法。

### 背景技术

[0002] 目标检测是人工智能的应用领域之一,在经过传统图像处理方法和卷积神经网络方法的发展后,深度学习方法展现出优异的效果,使目标检测得到了很好的提升。在使用深度学习方法进行目标检测的过程中,需要获取特征图,并根据其中的特征得到最终的目标位置和类别。另外,为了更广泛的实际应用,小目标检测是常见的任务,同时也成为对深度学习的挑战之一。在大小不同的目标检测中,对小目标的检测往往更为困难。与正常目标检测相比,小目标检测的困难点在于特征分辨率较小,导致最终类别检测结果的score相差不大且不高。由于小目标检测的困难点在于特征提取,目前的解决方法一般有改进特征提取网络,或使用tricks优化,或特征融合等其他方法。改进特征提取网络的方法固然有效,但改动较大且容易影响其他检测模块;使用tricks优化,例如soft-NMS通过后项处理方法优化检测结果,并未解决核心的困难点。而特征融合的方法即在特征层面进行改进,又不会过分影响网络结构,是常见的小目标检测解决方法。

[0003] 特征金字塔是常用的特征融合方法之一,通过特征金字塔可以获得不同层数融合的特征图。特征金字塔在应用于Faster R-CNN时,会根据RoI (Region of interest) 大小决定使用特征金字塔中不同层的输出。

[0004] 但是,这种方法仅仅利用了RoI大小的先验信息来决定使用特征金字塔不同的输出作为特征图,缺乏对后验信息的考虑,而且RoI的先验信息也只是属于预测性质。此外,并非所有的小目标均需要多层特征融合结果作为特征图才能有较好的检测结果,对大小相同的目标取相同层数融合的特征图进行处理,会产生冗余的消耗。

### 发明内容

[0005] 有鉴于此,本发明提供了一种自调节特征融合的小目标检测方法。

[0006] 为了实现上述目的,本发明采用如下技术方案:

[0007] 一种自调节特征融合的小目标检测方法,包括以下步骤:

[0008] 获取第一特征图,利用目标检测算法,计算第一目标回归结果和第一分类结果,并通过所述第一目标回归结果和所述第一分类结果,得到感兴趣区域信息和所述第一特征图的置信度score;所述感兴趣区域信息为感兴趣区域的高度和宽度;

[0009] 利用所述感兴趣区域的宽度和高度计算特征融合深度最大值;利用所述置信度score计算特征融合深度;

[0010] 根据所述特征融合深度最大值和所述特征融合深度利用特征金字塔获取特征融合后的第二特征图;

[0011] 将第二特征图利用所述目标检测算法,计算第二目标回归和第二分类结果,完成

目标检测。

[0012] 优选的,计算特征融合深度最大值具体如下:

[0013]  $d_{max} = \lfloor -\log_2(\sqrt{h*w}/N_i) \rfloor$  ;

[0014]  $N_i$ 表示标准图像输入边长, $h$ 是感兴趣区域的高度, $w$ 是感兴趣区域的宽度,且 $h, w \leq N_i$ 。

[0015] 优选的,计算特征融合深度具体如下:

[0016]  $d = \sum (D(S_1, S_i) \leq \theta \cdot \lambda)$  ;

[0017] 其中 $S_i$ 表示分类置信度score按从大到小顺序排列中第 $i$ 位的值, $D(\cdot)$ 表示 $S_1$ 与 $S_i$ 的距离,在此取 $D(S_1, S_i) = |S_1 - S_i|$ , $\lambda$ 表示阈值常量。

[0018] 优选的,特征金字塔获取特征图的具体过程如下:

[0019] 第一特征图输入到特征金字塔中,完成特征提取,各特征提取层输出不同的特征,经特征融合后输出多种特征图,根据所述特征融合深度最大值和所述特征融合深度选取第二特征图。

[0020] 经由上述的技术方案可知,与现有技术相比,本发明公开提供了一种自调节特征融合的小目标检测方法,通过利用预测RoI大小的先验信息结合score后验信息来调节特征融合深度,从而选择合适的特征融合结果作为特征图,综合考虑先验和后验信息可以提高特征获取的准确度;此外,通过类别score的差距影响特征融合深度,减少了对“大小较小但容易检测的目标”的冗余消耗。

## 附图说明

[0021] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

[0022] 图1附图为本发明的流程示意图;

[0023] 图2附图为本发明的特征金字塔图。

## 具体实施方式

[0024] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0025] 本发明实施例公开了一种自调节特征融合的小目标检测方法,如图1所示,包括以下步骤:

[0026] 获取第一特征图,利用目标检测算法,计算第一目标回归结果和第一分类结果,并通过第一目标回归结果和第一分类结果,得到感兴趣区域信息和第一特征图的置信度score;感兴趣区域信息为感兴趣区域的高度和宽度;

[0027] 利用感兴趣区域的宽度和高度计算特征融合深度最大值;利用置信度score计算特征融合深度;

[0028] 根据特征融合深度最大值和特征融合深度利用特征金字塔获取特征融合后的第二特征图；

[0029] 将第二特征图利用目标检测算法，计算第二目标回归和第二分类结果，完成目标检测。

[0030] 在本实施例中，具体实现流程包括以下步骤：

[0031] 在无特征融合条件下，计算目标回归和分类结果；

[0032] 例如Faster R-CNN算法模型，其特征图是取自最后一个卷积层的输出。在无特征融合条件下，正常计算该目标回归和分类结果，得到感兴趣区域的宽度和高度，以及其置信度score。

[0033] 这里的第一特征图可以为任意卷积神经网络的输出图像。

[0034] (2) 计算特征融合深度算法S2；

[0035] 算法包括如下两个具体步骤。

[0036] 步骤S21：计算特征融合深度最大值

$$[0037] \quad d_{\max} = \left\lfloor -\log_2 \left( \sqrt{h * w} / N_i \right) \right\rfloor \quad \textcircled{1}$$

[0038]  $N_i$ 表示标准图像输入边长，例如数据集为ImageNet时取 $N_i = 224$ 。 $h$ 和 $w$ 分别是RoI的高和宽，且 $h, w \leq N_i$ 。此外，令 $d_m$ 为特征可融合深度上限，例如在Faster R-CNN中 $d_m = 5$ 。通过①式计算 $d_{\max}$ ，若 $d_{\max} \leq d_m$ ，则 $d_{\max}$ 取①式；若 $d_{\max} > d_m$ ，则取 $d_{\max} = d_m$ 。以此限制保证最终结果 $d_{\max} \leq d_m$ 。

[0039] 步骤S22：计算特征融合深度

$$[0040] \quad d = \text{sum} (D(S_1, S_i) \leq \theta * \lambda) \quad \textcircled{2}$$

[0041] 其中 $S_i$ 表示分类置信度score按从大到小顺序排列中第 $i$ 位的值，例如 $S_1$ 表示score排名第1的类别所对应的值， $S_i$ 中的 $i = 2, 3, 4, 5$ 。 $D(\cdot)$ 表示 $S_1$ 与 $S_i$ 的距离，在此取 $D(S_1, S_i) =$

$$|S_1 - S_i| \cdot \lambda \text{ 表示阈值常量，可取 } \lambda = 0.4。 \theta \text{ 表示调整系数，取 } \theta = \sqrt{\frac{S_i}{\sum S_j}}, \text{ 其中 } j = 1, 2, 3, 4, 5。$$

$\text{sum}(\cdot)$ 表示计算符合不等式 $d(S_1, S_i) \leq \theta * \lambda$ 的个数。通过②式计算 $d$ ，若 $d \leq d_{\max}$ ，则 $d$ 取②式；若 $d > d_{\max}$ ，则取 $d = d_{\max}$ 。以此限制保证最终结果 $d \leq d_{\max}$ 。

[0042]  $d_m$ 和 $d_{\max}$ 决定了特征融合深度的上限，该上限表示特征金字塔自上而下特征融合的深度。②式中通过计算 $S_1$ 与其他 $S_i$ 的差距，反映排名第1的类别与其他类别的判断悬殊。由于score排名一般只有前几名的值较大，在此取排名前5的score值。若 $D(\cdot)$ 输出较小，表示预测两种类别的可能性相近，更容易满足不等式 $D(S_1, S_i) \leq \theta * \lambda$ 。 $\theta$ 表示 $S_i$ 的比重， $\theta$ 越大，越容易满足不等式 $D(S_1, S_i) \leq \theta * \lambda$ 。可见，只有当 $S_i$ 与 $S_1$ 相近且比重较大的 $S_i$ 才满足不等式 $D(S_1, S_i) \leq \theta * \lambda$ 。最后通过 $\text{sum}(\cdot)$ 计算特征融合深度。

[0043] (3) 根据(2)中的结果，获取特征融合后的特征图。

[0044] 如图2所示：

[0045] 特征金字塔的输入是标准图像。首先，标准图像输入到特征提取网络进行特征提取，各个特征提取层C1-C5输出不同的特征；然后，经过M2-M5的特征融合后，特征金字塔会输出P2-P6共5种深度的特征图，其中P6深度最小，P2深度最大；最后，根据(2)中计算得到的

特征融合深度 $d$ 来决定所采用的特征融合后的特征图,特征融合深度 $d$ 的值越大,则采用特征融合深度越大的特征图。例如特征融合深度 $d=1$ ,则output取P6输出;若特征融合深度 $d=2$ ,则output取P5输出,以此类推。output输出作为特征融合后的特征图。

[0046] (4) 在(3)中得到特征融合后的特征图以后,根据所使用的目标检测算法,利用特征图上的特征信息既可以回归计算目标的位置,又可以计算目标的类别置信度,因此可以通过计算得到最终目标回归和分类的结果,完成目标检测。例如所使用的目标检测算法是Faster R-CNN,首先将特征融合后的特征图输入至RPN网络得到RoI,接着将RoI结合特征融合后的特征图输入至分类回归网络,最后经过分类回归网络的计算,得到目标回归和分类的结果。

[0047] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

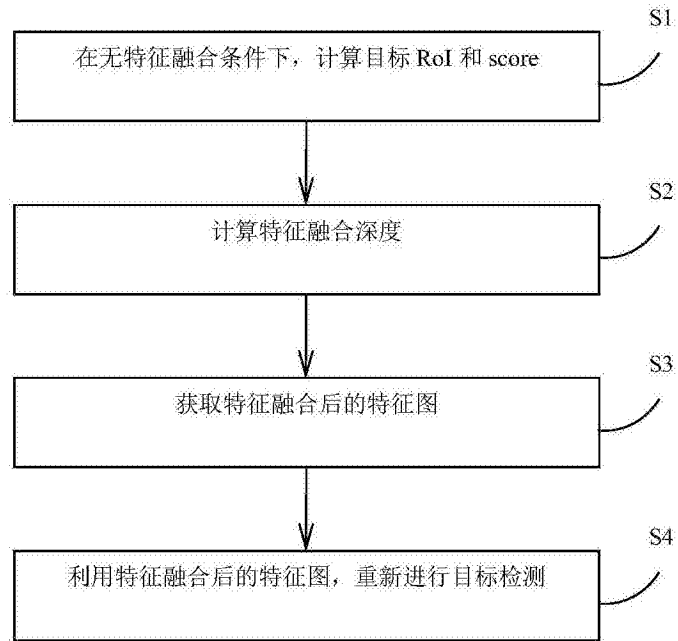


图1

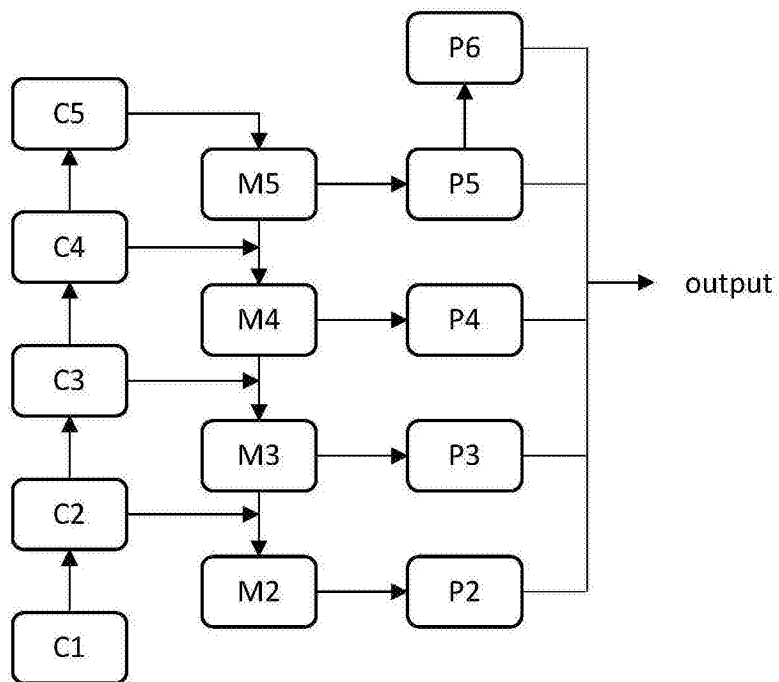


图2