



(12) 发明专利

(10) 授权公告号 CN 112561865 B

(45) 授权公告日 2024.03.12

(21) 申请号 202011406995.3

G06V 10/82 (2022.01)

(22) 申请日 2020.12.04

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 112561865 A

CN 110246571 A, 2019.09.17

CN 110349224 A, 2019.10.18

CN 110619350 A, 2019.12.27

(43) 申请公布日 2021.03.26

CN 110930421 A, 2020.03.27

CN 111882002 A, 2020.11.03

(73) 专利权人 深圳格瑞健康科技有限公司

地址 518000 广东省南山区南油华明路1号

后海居委会第二统建楼B座805

CN 111986217 A, 2020.11.24

US 2004029068 A1, 2004.02.12

US 2013158958 A1, 2013.06.20

(72) 发明人 黄少宏 赵志广 范卫华 李菊红

易超 林良强 李剑波 武剑

朱佳 刘勇 严志文 邢玉林

US 2018028294 A1, 2018.02.01

US 2019030371 A1, 2019.01.31

US 2019313963 A1, 2019.10.17

(74) 专利代理机构 广州嘉权专利商标事务所有

限公司 44205

专利代理师 常柯阳

段志伟. 基于特征融合和中心预测的目标检测算法. 现代计算机. 2020, (第09期), 全文.

王心醉;董宁宁;李欢利. 基于SIFT算法的图像特征点提取和匹配研究. 南京医科大学学报(自然科学版). 2013, (第02期), 全文.

审查员 何姣

(51) Int. Cl.

G06T 7/00 (2017.01)

G06V 10/25 (2022.01)

G06V 10/80 (2022.01)

权利要求书4页 说明书7页 附图1页

(54) 发明名称

恒磨牙位置的检测模型训练方法、系统和存储介质

(57) 摘要

本发明公开了一种恒磨牙位置的检测模型训练方法、系统和存储介质,方法包括以下步骤:获取若干张口腔牙列图像作为第一图像;对所述第一图像进行第一特征提取,得到第一特征图像;对所述第一特征图像在通道维度和空间维度分别进行选择性激活;在选择激活后的第一特征图像上进行第二特征提取;根据第二特征提取后的图像特征预测恒磨牙位置的检测框;对所述检测框的信息进行融合,生成第一置信度;将最大的第一置信度作为预测输出结果;根据所述预测输出结果反向更新所述检测模型的参数。本发明的检测模型在应用过程中更具有针对性且更能适应不同的应用情境,以提高检测模型在应用过

程中的检测结果的准确度。本发明可广泛应用于模型训练技术领域。



CN 112561865 B

1. 一种恒磨牙位置的检测模型训练方法,其特征在于,包括以下步骤:
 获取若干张口腔牙列图像作为第一图像;
 对所述第一图像进行第一特征提取,得到第一特征图像;
 对所述第一特征图像在通道维度和空间维度分别进行选择性的激活;
 在选择激活后的第一特征图像上进行第二特征提取;
 根据第二特征提取后的图像特征预测恒磨牙位置的检测框;
 对所述检测框的信息进行融合,生成第一置信度;
 将最大的第一置信度作为预测输出结果;
 根据所述预测输出结果反向更新所述检测模型的参数;
 其中,所述对所述第一特征图像在通道维度进行选择性的激活,包括:
 对所述第一特征图像进行平均池化和最大池化;
 采用注意力机制对平均池化和最大池化后的第一特征图像在通道维度进行选择性的激活;
 在通道维度得选择性激活中,采用公式1进行计算:

$$F^S = \sigma \left(\omega_1 \left(\omega_0 \left(F_{avg}^C \right) \right) + \omega_1 \left(\omega_0 \left(F_{max}^C \right) \right) \right) \text{公式1}$$

其中, F^S 为经过通道维度选择性激活后的特征, F_{avg}^C 表示经过平均池化后的图像特征, F_{max}^C 表示经过最大池化后的图像特征, ω_0 和 ω_1 为两个不同的矩阵参数用以表示针对不同通道数的权重大小, σ 表示激活函数sigmoid;

所述对所述第一特征图像在空间维度进行选择性的激活,其具体为:

采用卷积核对注意力机制处理后的第一图像在空间维度进行选择性的激活;
 在空间维度的选择性激活中,采用公式2进行计算:

$$F = \sigma \left(f^{7*7} \left(\left[F_{avg}^S; F_{max}^S \right] \right) \right) \text{公式2}$$

其中, f^{7*7} 表示7*7卷积运算, F_{avg}^S 表示在通道维度选择性激活所获得的特征基础上进行平均池化得到的图像特征, F_{max}^S 表示在通道维度选择性激活所获得的特征基础上进行最大池化得到的图像特征, σ 表示激活函数sigmoid;

所述对所述检测框的信息进行融合,生成第一置信度,包括:

将检测框的位置信息、第二置信度和图像特征信息进行融合;

根据融合结果生成第一置信度,所述第二置信度为图像本身的置信度,所述第一置信度为与所述第二置信度同一图像融合处理后的置信度;

其中,所述对所述检测框的信息进行融合时,在融合不同模态的信息时,采用公式3和公式4:

$$f_A = \sigma \left(W_{I_{gated}} f_{I_{embed}} + b_{I_{gated}} \right) f_{A_{embed}} + f_{A_{embed}} \text{公式3}$$

$$f_I = \sigma \left(W_{A_{gated}} f_{A_{embed}} + b_{A_{gated}} \right) f_{I_{embed}} + f_{I_{embed}} \text{公式4}$$

其中, f_A 为融合第二置信度信息后的图像特征, f_I 为融合图像特征后的置信度信息,

$f_{A_{embed}}$ 和 $f_{I_{embed}}$ 分别为原始的图像特征和置信度信息；

在融合不同模态的信息的基础上,结合图像的位置信息进行计算,以得到融合多模态信息的检测框置信度,其具体计算过程如公式5-公式8所示:

$$f_{fusion} = \sigma(W_S [f_A, f_I] + b_S) \text{ 公式5}$$

$$A := \text{soft max} \left(\left((1 \cdot p^T) \circ f_{fusion} U \right) V^T f_{G_{embed}} \right) \text{ 公式6}$$

$$f_{total,k} = \left(\left(f_{fusion} \right)^T U' \right)_K^T A \left(\left(f_{G_{embed}} \right)^T V' \right)_K \text{ 公式7}$$

$$f_S = \sigma(W_S f_{total}) + b_S \text{ 公式8}$$

其中, f_{fusion} 表示融合了第二置信度信息和图像特征信息后的特征, $f_{G_{embed}}$ 为编码后的检测框的位置信息, $f_{total,k}$ 则表示在 f_{fusion} 的基础上再进一步融合检测框的位置信息的特征, f_S 表示融合了多模态信息后最终得到的新置信度作为第一置信度。

2. 根据权利要求1所述的一种恒磨牙位置的检测模型训练方法,其特征在于,所述获取若干张口腔牙列图像作为第一图像,包括:

获取若干张口腔牙列图像;

对所述若干张口腔牙列图像的尺寸处理为预设尺寸大小;

对所述尺寸处理后的口腔牙列图像进行隔行采样,得到若干张子图像;

将所述若干张子图像进行拼接的图像作为第一图像。

3. 根据权利要求1所述的一种恒磨牙位置的检测模型训练方法,其特征在于,所述对所述第一图像进行第一特征提取,得到第一特征图像,包括:

对所述第一图像进行下采样特征提取,生成不同尺寸大小的特征图像;

将不同尺寸大小的特征图像作为第一特征图像。

4. 根据权利要求1所述的一种恒磨牙位置的检测模型训练方法,其特征在于,所述第二特征提取为抽象特征提取。

5. 一种恒磨牙位置的检测模型训练系统,其特征在于,包括:

获取模块,用于获取若干张口腔牙列图像作为第一图像;

第一特征提取模块,用于对所述第一图像进行第一特征提取,得到第一特征图像;

激活模块,用于对所述第一特征图像在通道维度和空间维度分别进行选择激活;

第二特征提取模块,用于在选择激活后的第一特征图像上进行第二特征提取;

预测模块,用于根据第二特征提取后的图像特征预测恒磨牙位置的检测框;

融合模块,用于对所述检测框的信息进行融合,生成第一置信度;

置信度选择模块,用于将最大的第一置信度作为预测输出结果;

参数更新模块,用于根据所述预测输出结果反向更新所述检测模型的参数;

其中,所述对所述第一特征图像在通道维度进行选择激活,包括:

对所述第一特征图像进行平均池化和最大池化;

采用注意力机制对平均池化和最大池化后的第一特征图像在通道维度进行选择激活;

在通道维度得选择性激活中,采用公式1进行计算:

$$F^S = \sigma \left(\omega_1 \left(\omega_0 \left(F_{avg}^C \right) \right) + \omega_1 \left(\omega_0 \left(F_{max}^C \right) \right) \right) \text{公式1}$$

其中, F^S 为经过通道维度选择性激活后的特征, F_{avg}^C 表示经过平均池化后的图像特征, F_{max}^C 表示经过最大池化后的图像特征, ω_0 和 ω_1 为两个不同的矩阵参数用以表示针对不同通道数的权重大小, σ 表示激活函数sigmoid;

所述对所述第一特征图像在空间维度进行选择性激活,其具体为:

采用卷积核对注意力机制处理后的第一图像在空间维度进行选择性激活;

在空间维度的选择性激活中,采用公式2进行计算:

$$F = \sigma \left(f^{7*7} \left(\left[F_{avg}^S; F_{max}^S \right] \right) \right) \text{公式2}$$

其中, f^{7*7} 表示7*7卷积运算, F_{avg}^S 表示在通道维度选择性激活所获得的特征基础上进行平均池化得到的图像特征, F_{max}^S 表示在通道维度选择性激活所获得的特征基础上进行最大池化得到的图像特征, σ 表示激活函数sigmoid;

所述对所述检测框的信息进行融合,生成第一置信度,包括:

将检测框的位置信息、第二置信度和图像特征信息进行融合;

根据融合结果生成第一置信度,所述第二置信度为图像本身的置信度,所述第一置信度为与所述第二置信度同一图像融合处理后的置信度;

其中,所述对所述检测框的信息进行融合时,在融合不同模态的信息时,采用公式3和公式4:

$$f_A = \sigma \left(W_{I_{gated}} f_{I_{embed}} + b_{I_{gated}} \right) f_{A_{embed}} + f_{A_{embed}} \text{公式3}$$

$$f_I = \sigma \left(W_{A_{gated}} f_{A_{embed}} + b_{A_{gated}} \right) f_{I_{embed}} + f_{I_{embed}} \text{公式4}$$

其中, f_A 为融合第二置信度信息后的图像特征, f_I 为融合图像特征后的置信度信息, $f_{A_{embed}}$ 和 $f_{I_{embed}}$ 分别为原始的图像特征和置信度信息;

在融合不同模态的信息的基础上,结合图像的位置信息进行计算,以得到融合多模态信息的检测框置信度,其具体计算过程如公式5-公式8所示:

$$f_{fusion} = \sigma \left(W_S [f_A, f_I] + b_S \right) \text{公式5}$$

$$A := \text{soft max} \left(\left((1 \cdot p^T) \circ f_{fusion} U \right) V^T f_{G_{embed}} \right) \text{公式6}$$

$$f_{total,k} = \left(\left(f_{fusion} \right)^T U' \right)_K^T A \left(\left(f_{G_{embed}} \right)^T V' \right)_K \text{公式7}$$

$$f_S = \sigma \left(W_S f_{total} \right) + b_S \text{公式8}$$

其中, f_{fusion} 表示融合了第二置信度信息和图像特征信息后的特征, $f_{G_{embed}}$ 为编码后的检测框的位置信息, $f_{total,k}$ 则表示在 f_{fusion} 的基础上再进一步融合检测框的位置信息的特征, f_S

表示融合了多模态信息后最终得到的新置信度作为第一置信度。

6. 一种恒磨牙位置的检测模型训练系统,其特征在於,包括:

至少一个存储器,用于存储程序;

至少一个处理器,用于加载所述程序以执行如权利要求1-4任一项所述的恒磨牙位置的检测模型训练方法。

7. 一种存储介质,其中存储有处理器可执行的程序,其特征在於,所述处理器可执行的程序在由处理器执行时用于执行如权利要求1-4任一项所述的恒磨牙位置的检测模型训练方法。

恒磨牙位置的检测模型训练方法、系统和存储介质

技术领域

[0001] 本发明涉及模型训练技术领域,尤其是一种恒磨牙位置的检测模型训练方法、系统和存储介质。

背景技术

[0002] 龋病是细菌(致龋菌)利用食物中的碳水化合物产酸导致牙齿硬组织进行性破坏的一种慢性疾病,是口腔主要的常见病,也是人类最普遍的疾病之一。6-12岁学龄儿童往往喜吃甜食、软食、黏食,容易粘在牙齿上,然而此阶段儿童口腔卫生习惯往往不佳或清洁方法掌握不好,很难能对牙齿进行有效的清洁,是患龋病的高发人群。除此之外,此阶段儿童处于替牙期,新萌出的恒牙尤其是新萌出的恒磨牙窝沟往往较深,容易积聚细菌且不易被清洁,细菌产生酸性分泌物会导致牙体硬组织破坏,极易产生龋齿。因此,对儿童口腔状况进行及时筛查并采取早期干预是预防恒牙龋病的治标之本。而窝沟封闭就是世界卫生组织推荐的预防恒牙龋坏的最佳方法,在许多国家和地区已普遍推广,我国也在全国各地大力推广学龄儿童窝沟封闭项目。窝沟封闭是指不损伤牙体组织,将窝沟封闭材料涂布于牙冠咬合面、颊舌面的窝沟点隙,当它流入并渗透入窝沟后固化变硬,形成一层保护性的屏障,覆盖在深窝沟上,能够阻止致龋菌及酸性代谢产物对牙体的侵蚀,以达到预防窝沟龋的方法。

[0003] 窝沟封闭的第一步是筛选出符合适应症的恒磨牙。以往开展窝沟封闭项目都是靠口腔医生人力筛查,需要花费大量人力、物力和财力。为了节省窝沟封闭项目的人力、物力和财力,提出了在线对牙齿进行检测和预警的方法。但是,由于目前的在线检测模型缺乏针对性和适应情境,导致在应用过程中,恒磨牙的识别结果准确度不高。

发明内容

[0004] 为解决上述技术问题之一,本发明的目的在于:提供一种恒磨牙位置的检测模型训练方法、系统和存储介质,其能提高检测模型在应用过程中检测结果的准确度。

[0005] 第一方面,本发明实施例提供了:

[0006] 一种恒磨牙位置的检测模型训练方法,包括以下步骤:

[0007] 获取若干张口腔牙列图像作为第一图像;

[0008] 对所述第一图像进行第一特征提取,得到第一特征图像;

[0009] 对所述第一特征图像在通道维度和空间维度分别进行选择激活;

[0010] 在选择激活后的第一特征图像上进行第二特征提取;

[0011] 根据第二特征提取后的图像特征预测恒磨牙位置的检测框;

[0012] 对所述检测框的信息进行融合,生成第一置信度;

[0013] 将最大的第一置信度作为预测输出结果;

[0014] 根据所述预测输出结果反向更新所述检测模型的参数。

[0015] 进一步地,所述获取若干张口腔牙列图像作为第一图像,包括:

- [0016] 获取若干张口腔牙列图像；
- [0017] 对所述若干张口腔牙列图像的尺寸处理为预设尺寸大小；
- [0018] 对所述尺寸处理后的口腔牙列图像进行隔行采样,得到若干张子图像；
- [0019] 将所述若干张子图像进行拼接的图像作为第一图像。
- [0020] 进一步地,所述对所述第一图像进行第一特征提取,得到第一特征图像,包括:
- [0021] 对所述第一图像进行下采样特征提取,生成不同尺寸大小的特征图像；
- [0022] 将不同尺寸大小的特征图像作为第一特征图像。
- [0023] 进一步地,所述对所述第一特征图像在通道维度进行选择激活,包括:
- [0024] 对所述第一特征图像进行平均池化和最大池化；
- [0025] 采用注意力机制对平均池化和最大池化后的第一特征图像在通道维度进行选择激活。进一步地,所述对所述第一特征图像在空间维度进行选择激活,其具体为:
- [0026] 采用卷积核对注意力机制处理后的第一图像在空间维度进行选择激活。
- [0027] 进一步地,所述第二特征提取为抽象特征提取。
- [0028] 进一步地,所述对所述检测框的信息进行融合,生成第一置信度,包括:
- [0029] 将检测框的位置信息、第二置信度和图像特征信息进行融合；
- [0030] 根据融合结果生成第一置信度。
- [0031] 第二方面,本发明实施例提供了:
- [0032] 一种恒磨牙位置的检测模型训练系统,包括:
- [0033] 获取模块,用于获取若干张口腔牙列图像作为第一图像；
- [0034] 第一特征提取模块,用于对所述第一图像进行第一特征提取,得到第一特征图像；
- [0035] 激活模块,用于对所述第一特征图像在通道维度和空间维度分别进行选择激活；
- [0036] 第二特征提取模块,用于在选择激活后的第一特征图像上进行第二特征提取；
- [0037] 预测模块,用于根据第二特征提取后的图像特征预测恒磨牙位置的检测框；
- [0038] 融合模块,用于对所述检测框的信息进行融合,生成第一置信度；
- [0039] 置信度选择模块,用于将最大的第一置信度作为预测输出结果；
- [0040] 参数更新模块,用于根据所述预测输出结果反向更新所述检测模型的参数。
- [0041] 第三方面,本发明实施例提供了:
- [0042] 一种恒磨牙位置的检测模型训练系统,包括:
- [0043] 至少一个存储器,用于存储程序；
- [0044] 至少一个处理器,用于加载所述程序以执行所述的恒磨牙位置的检测模型训练方法。
- [0045] 第四方面,本发明实施例提供了:
- [0046] 一种存储介质,其中存储有处理器可执行的程序,所述处理器可执行的程序在由处理器执行时用于执行所述的恒磨牙位置的检测模型训练方法。
- [0047] 本发明实施例的有益效果是:本发明实施例通过先对若干张口腔牙列图像进行第一特征提取,接着对第一特征图像在通道维度和空间维度分别进行选择激活,并在选择激活后的第一特征图像上进行第二特征提取,然后根据第二特征提取后的图像特征预测恒磨牙位置的检测框,并对检测框的信息进行融合,生成第一置信度,最后将最大的第一置信

度作为预测输出结果,根据预测输出结果反向更新所述检测模型的参数,从而使得通过本实施例训练后检测模型在应用过程中更具有针对性且更能适应不同的应用情境,以提高检测模型在应用过程中的检测结果的准确度。

附图说明

[0048] 图1为本发明一种具体实施例的恒磨牙位置的检测模型训练方法的流程图。

具体实施方式

[0049] 下面结合附图和具体实施例对本发明做进一步的详细说明。对于以下实施例中的步骤编号,其仅为了便于阐述说明而设置,对步骤之间的顺序不做任何限定,实施例中的各步骤的执行顺序均可根据本领域技术人员的理解来进行适应性调整。

[0050] 在以下的描述中,涉及到“一些实施例”,其描述了所有可能实施例的子集,但是可以理解,“一些实施例”可以是所有可能实施例的相同子集或不同子集,并且可以在不冲突的情况下相互结合。

[0051] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同。本文中所使用的术语只是为了描述本申请实施例的目的,不是旨在限制本申请。

[0052] 参照图1,本发明实施例提供了一种恒磨牙位置的检测模型训练方法,本实施例可应用于服务端或者各类平台的后台处理器,以使应用过程中,用户只需在人机交互界面进行指定步骤操作,既能完成恒磨牙位置检测。

[0053] 其中,本实施例包括以下步骤:

[0054] S1、获取若干张口腔牙列图像作为第一图像;在本步骤中,口腔牙列图像可以是实际检测过程中采集的被检测人员的图像,该采集过程可以是被采集人员通过手机或者其他具有摄像功能的设备拍摄得到。

[0055] 当该若干张口腔牙列图像应用于检测模型的训练过程中时,在一些实施例中,可采用数据翻转、旋转等数据增强方式对口腔牙列图像进行数据增广,以提供更多的训练样本数据,提高训练精度。

[0056] 在一些实施例中,所述步骤S1可通过以下方式实现:

[0057] 获取若干张口腔牙列图像;该若干张口腔牙列图像为不同的移动智能设备拍摄得到的图像,由于拍摄得到的图像尺寸不一致,而神经网络的输入尺寸是固定大小的,因此,对若干张口腔牙列图像的尺寸处理为预设尺寸大小。在这个过程中,可将图像尺寸设定为 $512*512$,即输入的图像数据大小为 $C*W*H$,其中, $C=3$ 表示输入的图像通道数为3,每个通道依次分别表示红、绿、蓝分量上的颜色值; $W=512$ 为图像的宽度; $H=512$ 为图像的高度。接着对尺寸处理后的口腔牙列图像进行隔行采样,得到若干张子图像;例如,对一张 $512*512$ 尺寸的图像,以图像左上角的一个 $2*2$ 区域中的每一个点为起点,按照长度为2的间隔对原始图像进行间隙采样,生成4张 $256*256$ 的子图像。最后,将若干张子图像进行拼接的图像作为检测模型输入的第一图像,以充分利用图像上的信息。

[0058] S2、对第一图像进行第一特征提取,得到第一特征图像。

[0059] 在本步骤中,是通过特征提取模块对第一图像进行第一特征提取。其中,特征提取

模块总由8层网络模块叠加组成,整体上这8层网络中2层一组组成一个下采样特征提取模块组,一共分为四组,每一组操作都将产生一张对应尺寸的特征图。通过下采样特征提取模块组进行特征提取后,可得到4张尺寸分别为512*512、256*256、128*128和64*64的特征图,最后将不同尺寸大小的特征图像作为第一特征图像。

[0060] S3、对第一特征图像在通道维度和空间维度分别进行选择性激活;

[0061] 在一些实施例中,对第一特征图像在通道维度进行选择性激活,其具有通过以下方式:

[0062] 对第一特征图像进行平均池化和最大池化;

[0063] 采用注意力机制对平均池化和最大池化后的第一特征图像在通道维度进行选择性激活。

[0064] 在一些实施例中,对第一特征图像在空间维度进行选择性激活,其可以采用以下方式:

[0065] 采用卷积核对注意力机制处理后的第一图像在空间维度进行选择性激活。

[0066] 在上述实施例中,注意力机制和卷积核组成注意力机制模块,具体地,首先进行通道维度的选择性激活,该部分将特征图分别经过最大池化和平均池化处理,随后将这两种操作得到的特征图分别选择性激活相应的特征通道,其中,激活通道的选择是一个可以通过神经网络迭代进行自学习的参数,最后两张分别在通道维度进行激活后得到的特征将会进行叠加融合,共同作为通道维度激活后的特征。其次是空间维度的选择性激活,该部分的是采用了同时利用平均池化及最大池化后的特征图作为输入的策略,并利用7*7的卷积核扩大感受野,将卷积后的输出通过sigmoid激活函数后映射到[0,1]内的数值表征特征图中各个位置相对于预测目标的重要程度。

[0067] 具体地,在通道维度的选择性激活中,采用公式1进行计算:

$$[0068] \quad F^S = \sigma \left(\omega_1 \left(\omega_0 \left(F_{avg}^C \right) \right) + \omega_1 \left(\omega_0 \left(F_{max}^C \right) \right) \right) \quad \text{公式 1}$$

[0069] 其中, F^S 为经过通道维度选择性激活后的特征, F_{avg}^C 表示经过平均池化后的图像特征, F_{max}^C 表示经过最大池化后的图像特征, ω_0 和 ω_1 为两个不同的矩阵参数用以表示针对不同通道数的权重大小, σ 表示激活函数sigmoid。

[0070] 在本步骤中,将两个相同的参数作用于不同方式处理后的特征图,可以使参数具备更强的鲁棒性,而连续运用两个参数对特征进行映射则能在保证特征图尺寸不变的基础上包含更多的神经元以拟合尽可能复杂的特征。最后,利用 σ 表示的一个激活函数完成从实域R到[0,1]的映射,以实现输出的归一化。

[0071] 在空间维度的选择性激活中,运用公式2进行计算:

$$[0072] \quad F = \sigma \left(f^{7*7} \left(\left[F_{avg}^S ; F_{max}^S \right] \right) \right) \quad \text{公式 2}$$

[0073] 其中, f^{7*7} 表示7*7卷积运算, F_{avg}^S 表示在通道维度选择性激活所获得的特征基础上进行平均池化得到的图像特征, F_{max}^S 表示在通道维度选择性激活所获得的特征基础上进行最大池化得到的图像特征, σ 表示激活函数sigmoid。

[0074] 在本步骤中,通过运用较大的卷积核来扩大感受野,从而帮助模型根据当前像素

点及其领域的内容来推断当前位置的重要程度。

[0075] S4、在选择激活后的第一特征图像上进行第二特征提取；在本步骤中，第二特征提取为抽象特征提取。

[0076] S5、根据第二特征提取后的图像特征预测恒磨牙位置的检测框；在本步骤中，其具体预测出检测框在第一图像上的位置和大小。

[0077] S6、对检测框的信息进行融合，生成第一置信度。

[0078] 在一些实施例中，步骤S6，其具体可通过以下方式实现：

[0079] 将检测框的位置信息、第二置信度和图像特征信息进行融合；

[0080] 根据融合结果生成第一置信度。

[0081] 在本实施例中，第二置信度为图像本身的置信度，第一置信度为与第二置信度同一图像融合处理后的置信度。

[0082] 由于现有的模型预测的检测框会存在一定的偏差，例如在同一侧中出现了两颗相同类型的牙齿，而这在实际情况中是不可能存在的现象。因此，在一些实施例中，通过采用多模态信息融合的检测框筛选机制。具体地，现有的检测框筛选方法仅仅是根据回归模型产生的置信度以及不同检测框间的重叠面积来判断，而忽略了候选检测框本身的图像特征信息。本实施例中候选框的位置信息和图像特征被视作两种不同模态的信息，将这两种不同模态的信息进行融合可以有效对检测框进行筛选。

[0083] 具体地，在融合不同模态的信息时，采用公式3和公式4：

$$[0084] \quad f_A = \sigma(W_{I_{gated}} f_{I_{embed}} + b_{I_{gated}}) f_{A_{embed}} + f_{A_{embed}} \quad \text{公式 3}$$

$$[0085] \quad f_I = \sigma(W_{A_{gated}} f_{A_{embed}} + b_{A_{gated}}) f_{I_{embed}} + f_{I_{embed}} \quad \text{公式 4}$$

[0086] 其中， f_A 为融合置信度信息后的图像特征， f_I 为融合图像特征后的置信度信息， $f_{A_{embed}}$ 和 $f_{I_{embed}}$ 分别为原始的图像特征和置信度信息，其余变量则为神经网络自学习的参数。

[0087] 在融合不同模态的信息的基础上，结合图像的位置信息进行计算，以得到融合多模态信息的检测框置信度，其具体计算过程如公式5-公式8所示：

$$[0088] \quad f_{fusion} = \sigma(W_S [f_A, f_I] + b_S) \quad \text{公式 5}$$

$$[0089] \quad A := \text{soft max} \left(\left((1 \cdot p^T) \circ f_{fusion} U \right) V^T f_{G_{embed}} \right) \quad \text{公式 6}$$

$$[0090] \quad f_{total,k} = \left(\left(f_{fusion} \right)^T U' \right)_K^T A \left(\left(f_{G_{embed}} \right)^T V' \right)_K \quad \text{公式 7}$$

$$[0091] \quad f_S = \sigma(W_S f_{total}) + b_S \quad \text{公式 8}$$

[0092] 其中， f_{fusion} 表示融合了置信度信息和图像特征信息后的特征， $f_{G_{embed}}$ 为编码后的检测框位置信息， $f_{total,k}$ 则表示在 f_{fusion} 的基础上再进一步融合检测框位置信息的特征， f_S 表示融合了多模态信息后最终得到的新置信度，其余字母或字母组合均表示神经网络所需要的参数。

[0093] S7、将最大的第一置信度作为预测输出结果；在本步骤中，是对每一个区域取置信度最大的作为预测输出结果。

[0094] S8、根据所述预测输出结果反向更新所述检测模型的参数,以使检测模型在应用过程中更具有针对性且更能适应不同的应用情境,以提高检测模型在应用过程中的检测结果的准确度。

[0095] 将上述实施例进行具体的应用,其包括以下步骤:

[0096] 采集移动智能设备采集的彩色图像,各个图像的角度、光线、所覆盖的区域情况均不相同。在本实施例中,图像的尺寸统一为512*512。

[0097] 将上述图像数据集被按照4:1的比例划分为了训练集和测试集,即数据集中的3316张图片用于训练模型,其余的829张图片用于验证模型的性能。最终的如表1所示的测试结果:

[0098] 表1

Method	AP	AP ₅₀	AP ₇₅	AR	AR ₅₀	SR ₇₅	AP _{molar1}	AP _{molar2}	Time (ms)
Baseline (D)	46.2	92.5	40.2	48.4	99.7	70.3	92.7	94.3	6.6
Baseline (O)	44.5	89.7	37.2	46.7	98.1	64.3	-	90.1	6.1
Baseline (A+D)	46.7	93.2	40.8	48.6	99.7	70.8	93.5	95.4	8.9
Baseline (AN+D)	47.9	94.5	41.6	48.5	98.9	72.5	94.2	97.3	10.4
Baseline (AN+A+D)	49.1	95.6	42.3	49.1	99.2	72.6	96.1	98.5	12.4

[0100] 在表1中,以AP和AR作为评估指标,即检测框与目标框的IOU值大于阈值时视为正确。其中,Baseline (D)表示同时检测第一前磨牙和第二前磨牙时的结果,Baseline (O)表示只检测第二前磨牙时的结果,A表示添加attention机制,AN表示添加A-NMS检测框筛选机制。

[0101] 综上所述,上述实施例通过增设注意力机制的方式将目标检测的范围进行限制,极大的增加了方法的适应性和鲁棒性。同时,设置检测框筛选机制,综合目标的图像特征和位置特征对检测框进一步的进行选择,保留图像特征与位置特征匹配的候选框,从而在相同数据集上提高检测结果的准确度。

[0102] 本发明实施例提供了一种与图1方法相对应的恒磨牙位置的检测模型训练系统,包括:

[0103] 获取模块,用于获取若干张口腔牙列图像作为第一图像;

[0104] 第一特征提取模块,用于对所述第一图像进行第一特征提取,得到第一特征图像;

[0105] 激活模块,用于对所述第一特征图像在通道维度和空间维度分别进行选择性激活;

[0106] 第二特征提取模块,用于在选择激活后的第一特征图像上进行第二特征提取;

[0107] 预测模块,用于根据第二特征提取后的图像特征预测恒磨牙位置的检测框;

[0108] 融合模块,用于对所述检测框的信息进行融合,生成第一置信度;

[0109] 置信度选择模块,用于将最大的第一置信度作为预测输出结果;

[0110] 参数更新模块,用于根据所述预测输出结果反向更新所述检测模型的参数。

[0111] 本发明方法实施例的内容均适用于本系统实施例,本系统实施例所具体实现的功能与上述方法实施例相同,并且达到的有益效果与上述方法达到的有益效果也相同。

[0112] 本发明实施例提供了一种恒磨牙位置的检测模型训练系统,包括:

[0113] 至少一个存储器,用于存储程序;

[0114] 至少一个处理器,用于加载所述程序以执行所述的恒磨牙位置的检测模型训练方法。

[0115] 本发明方法实施例的内容均适用于本系统实施例,本系统实施例所具体实现的功能与上述方法实施例相同,并且达到的有益效果与上述方法达到的有益效果也相同。

[0116] 本发明实施例提供了一种存储介质,其中存储有处理器可执行的程序,所述处理器可执行的程序在由处理器执行时用于执行所述的恒磨牙位置的检测模型训练方法。

[0117] 本发明实施例还公开了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存介质中。计算机设备的处理器可以从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行图1所示的方法。

[0118] 以上是对本发明的较佳实施进行了具体说明,但本发明并不限于所述实施例,熟悉本领域的技术人员在不违背本发明精神的前提下还可做作出种种的等同变形或替换,这些等同的变形或替换均包含在本申请权利要求所限定的范围内。



图1