



# (12) 发明专利

(10) 授权公告号 CN 111027589 B

(45) 授权公告日 2023. 04. 18

(21) 申请号 201911081589.1

G06V 10/774 (2022.01)

(22) 申请日 2019.11.07

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 111027589 A

CN 101038455 A, 2007.09.19

CN 101115131 A, 2008.01.30

CN 101489130 A, 2009.07.22

(43) 申请公布日 2020.04.17

CN 104867138 A, 2015.08.26

(73) 专利权人 成都傅立叶电子科技有限公司

CN 105787930 A, 2016.07.20

地址 610041 四川省成都市武侯区武侯新

CN 105956159 A, 2016.09.21

城管委会武兴二路17号

CN 106126872 A, 2016.11.16

专利权人 深圳市特发信息股份有限公司

CN 110070142 A, 2019.07.30

(72) 发明人 罗庚 陈英爽 蒋爽 徐涛

JP 2019185239 A, 2019.10.24

(74) 专利代理机构 成都诚中致达专利代理有限

US 2016366334 A1, 2016.12.15

公司 51280

US 2017237445 A1, 2017.08.17

专利代理师 阮涛 曹宇杰

WO 2016123815 A1, 2016.08.11

审查员 张玉晓

(51) Int. Cl.

G06V 10/776 (2022.01)

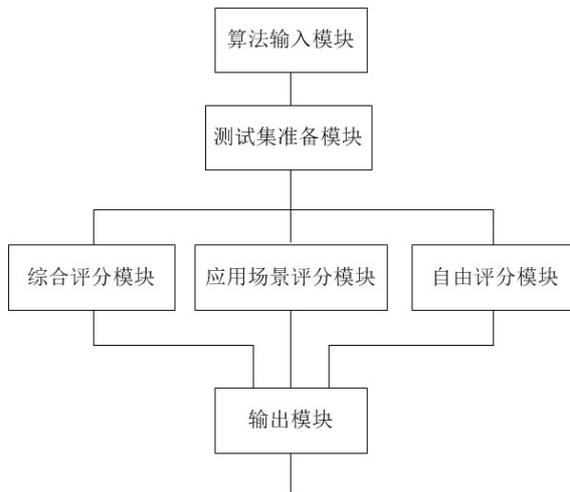
权利要求书3页 说明书10页 附图7页

(54) 发明名称

一种多分制目标检测算法评价系统及方法

(57) 摘要

多分制目标检测算法评价系统,包括算法输入模块,用于输入待测算法;测试集准备模块,根据待测算法选择具有被测目标种类的测试集;综合评分模块用于图像质量评分、图像分辨率评分、目标质量评分、目标大小评分加权得到总体综合评分;应用场景评分模块用于对场景图像分辨率评分、场景目标质量评分、场景目标大小评分加权得到应用场景评分;自由评分模块,根据用户测试需求自定义参数生成对应测试集以完成针对性的测评。同时公开一种多分制目标检测算法评价方法。能给出总体综合评分、应用场景评分、综合自由评分,不仅能客观反映出算法的好坏,还能反映出算法的好坏领域,而且能给算法设计者明确的优化方向,使算法优化更有针对性和目的性。



1. 一种多分制目标检测算法评价系统,其特征在于,包括:

算法输入模块,用于输入待测算法;

测试集准备模块,用于根据待测算法选择具有被测目标种类的测试集;

综合评分模块,包括:

总体图像质量数据集生成单元,用于对测试集进行加噪处理,并根据图像质量评价指标获得测试集每张图像的距离指标,根据距离指标将测试集划分为多个质量等级的数据集;

总体图像分辨率数据集生成单元,用于根据图像分辨率大小将测试集划分为多个不同等级的数据集;

总体目标质量数据集生成单元,用于根据目标质量好差将测试集划分为多个不同等级的数据集;

总体目标大小数据集生成单元,用于根据目标大小将测试集划分为多个不同等级的数据集;

总体图像质量评分单元,用于通过总体图像质量数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用归一化权重对各等级对应的mAP评分进行加权处理获得总体图像质量评分;

总体图像分辨率评分单元,用于通过总体图像分辨率数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用平均权重对各等级对应的mAP评分进行加权处理获得总体图像分辨率评分;

总体目标质量评分单元,用于通过总体目标质量数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用归一化权重对各等级对应的mAP评分进行加权处理获得总体目标质量评分;

总体目标大小评分单元,用于通过总体目标数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用平均权重对各等级对应的mAP评分进行加权处理获得总体目标大小评分;

总体综合评分单元,用于对总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分进行平均加权处理获得总体综合评分;

应用场景评分模块,包括:

场景图像质量数据集生成单元,用于对测试集进行加噪处理,并根据图像质量评价指标获得测试集每张图像的距离指标,根据距离指标将测试集划分为多个质量等级的数据集,将其中等级最好的两个数据集中的数据集作为场景图像质量数据集;

场景图像分辨率数据集生成单元,用于根据预设的分辨率参数从测试集中生成与该预设的分辨率参数对应的场景图像分辨率数据集;

场景目标质量数据集生成单元,用于根据预设的目标质量参数从测试集中生成与该预设的目标质量参数对应的场景目标质量数据集;

场景目标大小数据集生成单元,用于根据预设的目标大小参数从测试集中生成与该预设的目标大小参数对应的场景目标大小数据集;

场景图像质量评分单元,用于通过场景图像质量数据集生成单元生成的场景图像质量数据集对待测算法进行测试,计算mAP以获得场景图像质量评分;

场景图像分辨率评分单元,用于通过场景图像分辨率数据集生成单元生成的数据集对待测算法进行测试,计算mAP以获得场景图像分辨率评分;

场景目标质量评分单元,用于通过场景目标质量数据集生成单元生成的数据集对待测算法进行测试,计算mAP以获得场景目标质量评分;

场景目标大小评分单元,用于通过场景目标大小数据集生成单元生成的数据集对待测算法进行测试,计算mAP以获得场景目标大小评分;

场景综合评分单元,用于对场景图像质量评分、场景图像分辨率评分、场景目标质量评分、场景目标大小评分进行平均加权处理获得应用场景评分;

自由评分模块,包括:

自由场景预设单元,用于响应于用户输入的图像质量参数、图像分辨率参数、目标质量参数、目标大小参数中至少一种自由场景参数,对应选择出自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中与自由场景参数对应的场景作为自由测试场景;

自由场景数据集生成单元,根据自由测试场景及自由场景参数从测试集中生成对应的自由数据集;

自由评分单元,用于通过自由数据集对待测算法进行测试,计算mAP以获得对应自由测试场景的自由评分,并在获得两种以上自由评分时,将各自由评分进行平均加权得到综合自由评分;在只获得一种自由评分时,以该种自由评分作为综合自由评分。

2. 根据权利要求1所述的多分制目标检测算法评价系统,其特征在于,综合评分模块还包括归一化权重生成单元,用于对不同等级mAP评分设定不同等级的同程度分;根据不同等级取得同程度分时分别对应需要的mAP分数,计算各等级对应的权重;对各等级对应的权重进行归一化处理,获得各等级对应的归一化权重。

3. 根据权利要求1所述的多分制目标检测算法评价系统,其特征在于,应用场景评分模块还包括场景预设单元,用于响应于用户的输入以完成分辨率参数、目标质量参数、目标大小参数的预设。

4. 根据权利要求1所述的多分制目标检测算法评价系统,其特征在于,还包括:输出模块,用于输出评价结果并予以显示和/或输出包含评价结果的可打印文档。

5. 根据权利要求4所述的多分制目标检测算法评价系统,其特征在于,评价结果为总体综合评分、应用场景评分、综合自由评分中至少一种。

6. 一种多分制目标检测算法评价方法,其特征在于,包括步骤:

根据输入的待测算法选择具有被测目标种类的测试集;

响应于用户的输入选择,执行综合评分方法、应用场景评分方法、自由评分方法中至少一种评分方法;其中:

所述综合评分方法,包括步骤:

对测试集进行加噪处理,并根据图像质量评价指标获得测试集每张图像的距离指标,根据距离指标将测试集划分为多个质量等级的数据集,并分别对待测算法进行测试,分别计算mAP,利用归一化权重对各等级对应的mAP评分进行加权处理获得总体图像质量评分;

根据图像分辨率大小将测试集划分为多个不同等级的数据集,并分别对待测算法进行测试,分别计算mAP,利用平均权重对各等级对应的mAP评分进行加权处理获得总体图像分

辨率评分；

根据目标质量好差将测试集划分为多个不同等级的数据集，并分别对待测算法进行测试，分别计算mAP，利用归一化权重对各等级对应的mAP评分进行加权处理获得总体目标质量评分；

根据目标大小将测试集划分为多个不同等级的数据集，并分别对待测算法进行测试，分别计算mAP，利用平均权重对各等级对应的mAP评分进行加权处理获得总体目标大小评分；

对总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分进行平均加权处理获得总体综合评分；

所述应用场景评分方法，包括步骤：

对测试集进行加噪处理，并根据图像质量评价指标获得测试集每张图像的距离指标，根据距离指标将测试集划分为多个质量等级的数据集，将其中等级最好的两个数据集中的数据集作为场景图像质量数据集；通过场景图像质量数据集对待测算法进行测试，计算mAP以获得场景图像质量评分；

根据预设的分辨率参数从测试集中生成与该预设的分辨率参数对应的场景图像分辨率数据集；通过场景图像分辨率数据集对待测算法进行测试，计算mAP以获得场景图像分辨率评分；

根据预设的目标质量参数从测试集中生成与该预设的目标质量参数对应的场景目标质量数据集；通过场景目标质量数据集对待测算法进行测试，计算mAP以获得场景目标质量评分；

根据预设的目标大小参数从测试集中生成与该预设的目标大小参数对应的场景目标大小数据集；通过场景目标大小数据集对待测算法进行测试，计算mAP以获得场景目标大小评分；

对场景图像质量评分、场景图像分辨率评分、场景目标质量评分、场景目标大小评分进行平均加权处理获得应用场景评分；

所述自由评分方法，包括步骤：

响应于用户输入的图像质量参数、图像分辨率参数、目标质量参数、目标大小参数中至少一种自由场景参数，对应选择出自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中自由场景参数对应的场景作为自由测试场景；

根据自由测试场景及自由场景参数从测试集中生成对应的自由数据集；

通过自由数据集对待测算法进行测试，计算mAP以获得对应自由测试场景的自由评分，并在获得两种以上自由评分时，将各自由评分进行平均加权得到综合自由评分；在只获得一种自由评分时，以该种自由评分作为综合自由评分。

## 一种多分制目标检测算法评价系统及方法

### 技术领域

[0001] 本发明涉及图像处理与目标检测技术,尤其与一种多分制目标检测算法评价系统及方法相关。

### 背景技术

[0002] 目标检测技术是近年来计算机视觉领域中备受关注的方向,主要涉及计算机视觉、图像处理、人工智能、模式识别等学科,被广泛地应用于靶场测量、目标监控、视频压缩、车辆跟踪和航空航天等方面。

[0003] 目前,在对目标检测算法的测评中,通常用准确率和召回率来描述一种算法的好坏,而mAP(mean Average Precision)及F-Score作为一个兼顾考虑以上两种指标,并被广泛应用。然而,传统的mAP或F-Score计算方法并不能定量反映出算法在哪些领域的擅长之处或弱点,也不能为算法设计者提供优化点,且通常存在评价片面、针对性不够强不够全面,也不能够根据应用场景进行倾向性评价。因此,需要开发多分制的评价系统,结合多种评分参数及多种评价类型,提供更加全面的评价功能。

### 发明内容

[0004] 本发明主要针对相关现有技术的不足,提供一种多分制目标检测算法评价系统及方法,能给出总体综合评分、应用场景评分、综合自由评分,不仅能客观反映出算法的好坏,还能反映出算法的好坏领域,而且能给算法设计者明确的优化方向,使算法优化更有针对性和目的性。

[0005] 为了实现上述目的,本发明采用以下技术:

[0006] 一种多分制目标检测算法评价系统,其特征在于,包括:

[0007] 算法输入模块,用于输入待测算法;

[0008] 测试集准备模块,用于根据待测算法选择具有被测目标种类的测试集;

[0009] 综合评分模块,用于根据图像质量、图像分辨率、目标质量、目标大小将测试集分别划分为四类具有多个不同等级的数据集,其中,在根据图像质量将测试集划分前,先对测试集进行加噪处理;通过对各不同类的数据集分别计算mAP以分别获得总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分,并将各总体评分进行平均加权得到总体综合评分;

[0010] 应用场景评分模块,用于对测试集进行加噪处理并根据图像质量将测试集划分为不同等级的数据集,以等级最好的两个数据集中的数据集计算mAP以获得场景图像质量评分,并根据预设的图像分辨率参数、目标质量参数、目标大小参数分别从测试集中分别生成对应的场景数据集,以不同的场景数据集分别计算mAP以分别获得场景图像分辨率评分、场景目标质量评分、场景目标大小评分,并将各场景评分进行平均加权得到应用场景评分;

[0011] 自由评分模块,用于根据预设的自由测试场景及对应预设的自由场景参数从测试

集中生成对应的自由数据集,其中,预设的自由测试场景包括自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中至少一种;并通过自由数据集计算mAP 以获得对应自由测试场景的自由评分,且在获得两种以上自由评分时,将各自由评分进行平均加权得到综合自由评分;在只获得一种自由评分时,以该种自由评分作为综合自由评分。

[0012] 一种多分制目标检测算法评价方法,其特征在于,包括步骤:

[0013] 根据输入的待测算法选择具有被测目标种类的测试集;

[0014] 响应于用户的输入选择,执行综合评分方法、应用场景评分方法、自由评分方法中至少一种评分方法;其中:

[0015] 综合评分方法,包括步骤:

[0016] 根据图像质量、图像分辨率、目标质量、目标大小将测试集分别划分为四类具有多个不同等级的数据集,其中,在根据图像质量将测试集划分前,先对测试集进行加噪处理;

[0017] 通过对各不同类的数据集分别计算mAP以分别获得总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分;

[0018] 将各总体评分进行平均加权得到总体综合评分;

[0019] 应用场景评分方法,包括步骤:

[0020] 对测试集进行加噪处理并根据图像质量将测试集划分为不同等级的数据集,以等级最好的两个数据集中的数据集计算mAP以获得场景图像质量评分;

[0021] 根据预设的图像分辨率参数、目标质量参数、目标大小参数分别从测试集中分别生成对应的场景数据集;

[0022] 以不同的场景数据集分别计算mAP以分别获得场景图像分辨率评分、场景目标质量评分、场景目标大小评分;

[0023] 将各场景评分进行平均加权得到应用场景评分;

[0024] 自由评分方法,包括步骤:

[0025] 根据预设的自由测试场景及对应预设的自由场景参数从测试集中生成对应的自由数据集,其中,预设的自由测试场景包括自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中至少一种;

[0026] 通过自由数据集计算mAP 以获得对应自由测试场景的自由评分;

[0027] 在获得两种以上自由评分时,将各自由评分进行平均加权得到综合自由评分;在只获得一种自由评分时,以该种自由评分作为综合自由评分。

[0028] 本发明有益效果在于:

[0029] 1、传统评分方法是对一批测试集直接进行目标检测,得到预测标签,并采用mAP或FnScore的方法得到预测标签与真实标签之间的评分,由于测试集没做任务限定,因此不能定量反映出算法在哪些领域的擅长之处或弱点,也不能与算法设计者提供优化点;相比于传统方法,本发明的方法/系统,能根据评价需求给出总体综合评分、应用场景评分、综合自由评分,具体到综合图像质量评分、综合图像分辨率评分、综合图像目标质量评分、综合图像目标大小评分;以及应用场景图像质量评分、场景图像分辨率评分、场景图像目标质量评分、应用场景图像目标大小评分,还可以进行自由的评分,不仅能客观反映出算法的好坏,还能反映出算法的好坏领域,而且能给算法设计者明确的优化方向,使算法优化更有针对

性和目的性；

[0030] 2、总体综合评分可以客观反映待测算法的总体综合评分，具体的，在本发明的总体综合评分中同时结合了图像质量、图像分辨率、目标质量、目标大小进行综合总评，其中，图像质量和目标质量评分中采用了归一化权重处理方式，目标质量和目标大小采用平均权重处理方式，总评采用平均权重处理方式，更加全面的考虑及结合了不同图像属性的因素，利于对待测算法进行更全面、有效、准确的测评。

[0031] 3、通过应用场景评分，可以实现专家/用户参数的设置，针对性的评价待测算法在特定应用场景时的效果；

[0032] 4、通过自由评分，可以根据某些算法对某些特定场景项目进行单独测评，比如只对算法进行目标大小的评分，或分辨率大小评分，又或以上两项或其他项目进行评分，使用者可按照自己的想法，自由评分。

## 附图说明

[0033] 图1为本申请实施例的评价系统结构框图。

[0034] 图2为本申请实施例的综合评分模块结构框图。

[0035] 图3为本申请实施例的应用场景评分模块结构框图。

[0036] 图4为本申请实施例的自由评分模块结构框图。

[0037] 图5为本申请实施例的评价方法步骤流程图。

[0038] 图6为本申请实施例的综合评分法步骤流程图。

[0039] 图7为本申请实施例的应用场景评分法步骤流程图。

[0040] 图8为本申请实施例的自由评分法步骤流程图。

## 具体实施方式

[0041] 下面结合附图，详细说明本申请的系统及方法的实施手段/方式，并附实例。

[0042] 具体的，作为本申请的实施例，评价系统的整体结构框架如图1所示：

[0043] 系统包括算法输入模块、测试集准备模块、综合评分模块、应用场景评分模块、自由评分模块，以及输出模块。

[0044] 算法输入模块连接测试集准备模块，测试集准备模块分别连接综合评分模块、应用场景评分模块、自由评分模块，综合评分模块、应用场景评分模块、自由评分模块分别连接输出模块。

[0045] 具体的：

[0046] 算法输入模块，作为发起端，用于供用户输入待测算法，并将待测算法向后级模块传输。

[0047] 测试集准备模块，用于根据待测算法选择具有被测目标种类的测试集。

[0048] 输出模块，用于输出评价结果并予以显示和/或输出包含评价结果的可打印文档。

[0049] 评价结果包括有总体综合评分、应用场景评分、综合自由评分，根据用户的操作需求，输出至少一种，也可以输出多种或者全部。

[0050] 下面对三大评分模块做详细的说明：

[0051] (1)综合评分模块，用于根据图像质量、图像分辨率、目标质量、目标大小将测试集

分别划分为四类具有多个不同等级的数据集,其中,在根据图像质量将测试集划分前,先对测试集进行加噪处理;通过对各不同类的数据集分别计算mAP以分别获得总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分,并将各总体评分进行平均加权得到总体综合评分。

[0052] 具体的,如图2所示,综合评分模块包括:

[0053] 4个数据集生成单元:总体图像质量数据集生成单元、总体图像分辨率数据集生成单元、总体目标质量数据集生成单元、总体目标大小数据集生成单元;

[0054] 5个评分单元:总体图像质量评分单元、总体图像分辨率评分单元、总体目标质量评分单元、总体目标大小评分单元,以及总体综合评分单元;

[0055] 1个归一化权重生成单元,为总体图像质量评分单元和总体目标质量评分单元提供归一化权重。具体的是:对不同等级mAP评分设定不同等级的同程度分;根据不同等级取得同程度分时分别对应需要的mAP分数,计算各等级对应的权重;对各等级对应的权重进行归一化处理,获得各等级对应的归一化权重。

[0056] 如图2所示,其中的数据集生成单元分别连接对应的评分单元,各分类评分单元分别连接至总体综合评分单元。归一化权重生成单元,分别连接总体图像质量评分单元和总体目标质量评分单元。

[0057] 总体图像质量评分:

[0058] 首先总体图像质量数据集生成单元,对测试集进行加噪处理,主要是对测试集的图像加入不同程度的高斯噪声、椒盐噪声或其他噪声,和/或,进行不同程度的平滑处理,扩大测试集图像数量,并区分出不同程度噪声等级的图集。并采用SSIM、MS-SSIM、IW-SSIM、FSIM、MDSI中一种或多种结合作为图像质量指标获得测试集每张图像的距离指标,根据距离指标将测试集划分为多个质量等级的数据集。

[0059] 然后,总体图像质量评分单元通过总体图像质量数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用归一化权重对各等级对应的mAP评分进行加权处理获得总体图像质量评分。

[0060] 总体图像分辨率评分:

[0061] 首先,总体图像分辨率数据集生成单元,根据图像分辨率大小将测试集划分为多个不同等级的数据集。

[0062] 然后,总体图像分辨率评分单元,通过总体图像分辨率数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用平均权重对各等级对应的mAP评分进行加权处理获得总体图像分辨率评分。

[0063] 总体目标质量评分:

[0064] 首先,总体目标质量数据集生成单元,根据目标质量好差将测试集划分为多个不同等级的数据集。

[0065] 然后,总体目标质量评分单元,通过总体目标质量数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用归一化权重对各等级对应的mAP评分进行加权处理获得总体目标质量评分。

[0066] 总体目标大小评分:

[0067] 首先,总体目标大小数据集生成单元,根据目标大小将测试集划分为多个不同等

级的数据集。

[0068] 然后,总体目标大小评分单元,通过总体目标数据集生成单元生成的数据集分别对待测算法进行测试,分别计算mAP,利用平均权重对各等级对应的mAP评分进行加权处理获得总体目标大小评分。

[0069] 总体综合评分:

[0070] 由总体综合评分单元,对总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分进行平均加权处理获得总体综合评分。

[0071] (3)应用场景评分模块,用于对测试集进行加噪处理并根据图像质量将测试集划分为不同等级的数据集,以等级最好的两个数据集中的数据集计算mAP以获得场景图像质量评分,并根据预设的图像分辨率参数、目标质量参数、目标大小参数分别从测试集中分别生成对应的场景数据集,以不同的场景数据集分别计算mAP以分别获得场景图像分辨率评分、场景目标质量评分、场景目标大小评分,并将各场景评分进行平均加权得到应用场景评分。

[0072] 具体的,如图3所示,应用场景评分模块包括:

[0073] 1个预设单元:场景预设单元,用于响应于用户的输入以完成分辨率参数、目标质量参数、目标大小参数的预设;

[0074] 4个数据集生成单元:场景图像质量数据集生成单元、场景图像分辨率数据集生成单元、场景目标质量数据集生成单元、场景目标大小数据集生成单元;

[0075] 5个评分单元:场景图像质量评分单元、场景图像分辨率评分单元、场景目标质量评分单元、场景目标大小评分单元,以及场景综合评分单元。

[0076] 如图3所示,场景预设单元连接场景图像分辨率数据集生成单元、场景目标质量数据集生成单元、场景目标大小数据集生成单元,用于将用户输入的对应参数,提供给对应的数据生成单元。场景图像质量数据集生成单元连接场景图像质量评分单元,场景图像分辨率数据集生成单元连接场景图像分辨率评分单元,场景目标质量数据集生成单元连接场景目标质量评分单元,场景目标大小数据集生成单元连接场景目标大小评分单元。场景综合评分单元分别连接场景图像质量评分单元、场景图像分辨率评分单元、场景目标质量评分单元、场景目标大小评分单元。

[0077] 场景图像质量评分:

[0078] 首先,场景图像质量数据集生成单元,对测试集进行加噪处理,并根据图像质量评价指标获得测试集每张图像的距离指标,根据距离指标将测试集划分为多个质量等级的数据集,将其中等级最好的两个数据集中的数据集作为场景图像质量数据集。

[0079] 然后,场景图像质量评分单元,通过场景图像质量数据集生成单元生成的场景图像质量数据集对待测算法进行测试,计算mAP以获得场景图像质量评分。

[0080] 场景图像分辨率评分:

[0081] 首先,用户通过场景预设单元输入图像分辨率参数,场景预设单元响应于用户的输入以完成分辨率参数的预设。

[0082] 然后,场景图像分辨率数据集生成单元,根据预设的分辨率参数从测试集中生成与该预设的分辨率参数对应的场景图像分辨率数据集。

[0083] 再后,场景图像分辨率评分单元,通过场景图像分辨率数据集生成单元生成的数

据集对待测算法进行测试,计算mAP以获得场景图像分辨率评分。

[0084] 场景目标质量评分:

[0085] 首先,用户通过场景预设单元输入目标质量参数,场景预设单元响应于用户的输入以完成目标质量参数的预设。

[0086] 然后,场景目标质量数据集生成单元,根据预设的目标质量参数从测试集中生成与该预设的目标质量参数对应的场景目标质量数据集。

[0087] 再后,场景目标质量评分单元,通过场景目标质量数据集生成单元生成的数据集对待测算法进行测试,计算mAP以获得场景目标质量评分。

[0088] 场景目标大小评分:

[0089] 首先,用户通过场景预设单元输入目标大小参数,场景预设单元响应于用户的输入以完成目标大小参数的预设。

[0090] 然后,场景目标大小数据集生成单元,根据预设的目标大小参数从测试集中生成与该预设的目标大小参数对应的场景目标大小数据集。

[0091] 再后,场景目标大小评分单元,通过场景目标大小数据集生成单元生成的数据集对待测算法进行测试,计算mAP以获得场景目标大小评分。

[0092] 场景综合评分:

[0093] 场景综合评分单元,对场景图像质量评分、场景图像分辨率评分、场景目标质量评分、场景目标大小评分进行平均加权处理获得应用场景评分。

[0094] (3)自由评分模块,用于根据预设的自由测试场景及对应预设的自由场景参数从测试集中生成对应的自由数据集,其中,预设的自由测试场景包括自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中至少一种;并通过自由数据集计算mAP 以获得对应自由测试场景的自由评分,且在获得两种以上自由评分时,将各自由评分进行平均加权得到综合自由评分;在只获得一种自由评分时,以该种自由评分作为综合自由评分。

[0095] 如图4所示,自由评分模块包括依次连接的自由场景预设单元、自由场景数据集生成单元、自由评分单元。

[0096] 在需要进行自由评分时:

[0097] 首先,用户根据自己的需求输入图像质量参数、图像分辨率参数、目标质量参数、目标大小参数中至少一种自由场景参数,自由场景预设单元响应到用户的输入,对应选择出自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中与用户输入的自由场景参数对应的场景作为自由测试场景。

[0098] 然后,自由场景数据集生成单元,根据自由测试场景及自由场景参数从测试集中生成对应的自由数据集。

[0099] 再后,自由评分单元,通过自由数据集对待测算法进行测试,计算mAP以获得对应自由测试场景的自由评分。其中:

[0100] 在获得两种以上自由评分时,将各自由评分进行平均加权得到综合自由评分;

[0101] 在只获得一种自由评分时,以该种自由评分作为综合自由评分。

[0102] 如图5所示,本申请实施例的评价方法步骤流程图。

[0103] 一种多分制目标检测算法评价方法,包括步骤:

[0104] (1) 根据输入的待测算法选择具有被测目标种类的测试集。

[0105] (2) 响应于用户的输入选择, 执行综合评分方法、应用场景评分方法、自由评分方法中至少一种评分方法。

[0106] 具体的, 综合评分方法的详细步骤如图6所示:

[0107] (1) 根据图像质量、图像分辨率、目标质量、目标大小将测试集分别划分为四类具有多个不同等级的数据集, 其中, 在根据图像质量将测试集划分前, 先对测试集进行加噪处理。

[0108] (2) 通过对各不同类的数据集分别计算mAP以分别获得总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分。

[0109] (3) 将各总体评分进行平均加权得到总体综合评分。

[0110] 具体的, 应用场景评分方法的详细步骤如图7所示:

[0111] (1) 对测试集进行加噪处理并根据图像质量将测试集划分为不同等级的数据集, 以等级最好的两个数据集中的数据集计算mAP以获得场景图像质量评分。

[0112] (2) 根据预设的图像分辨率参数、目标质量参数、目标大小参数分别从测试集中分别生成对应的场景数据集。

[0113] (3) 以不同的场景数据集分别计算mAP以分别获得场景图像分辨率评分、场景目标质量评分、场景目标大小评分。

[0114] (4) 将各场景评分进行平均加权得到应用场景评分。

[0115] 具体的, 自由评分方法的详细步骤如图8所示:

[0116] (1) 根据预设的自由测试场景及对应预设的自由场景参数从测试集中生成对应的自由数据集, 其中, 预设的自由测试场景包括自由图像质量场景、自由图像分辨率场景、自由目标质量场景、自由目标大小场景中至少一种。

[0117] (2) 通过自由数据集计算mAP 以获得对应自由测试场景的自由评分:

[0118] (2-1) 在获得两种以上自由评分时, 将各自由评分进行平均加权得到综合自由评分;

[0119] (2-2) 在只获得一种自由评分时, 以该种自由评分作为综合自由评分。

[0120] 下面进一步举例说明本申请的系统及方法。

[0121] 通过输入模块, 输入待测算法。

[0122] 测试集准备模块, 在总体测试集中, 选择具有被测目标种类的原测试集, 命名为测试集TG。

[0123] 根据用户的测试/测评需求, 选择综合评分、应用场景评分、自由评分中至少一种进行测评。选择综合评分, 无需用户设定参数。选择应用场景评分, 用户需要通过场景预设单元输入对应场景测评需要的分辨率参数、目标质量参数、目标大小参数。选择自由评分, 用户可以自定义输入需要测试的属性, 以及需要测评的参数限定, 如需要输入图像质量参数、图像分辨率参数、目标质量参数、目标大小参数中至少一种, 自由场景预设单元根据输入有值且值不为零的对应参数, 确定需要测评的应用场景, 若用户只输入了一种参数, 则只确定一种对应的场景, 只完成对应场景参数下的测评。

[0124] 一、若选择需要进行综合评分, 则综合评分模块运作, 开始进行以下功能/步骤:

[0125] (1) 总体图像质量评分:

[0126] 对TG测试集进行加噪处理,分别提取5种质量等级的5个测试集T1~T5。

[0127] 计算T1与P1的mAP1, 计算T2与P2的mAP2,同理得到mAP3、mAP4、mAP5,因此,得出5个等级的评分,并利用归一化权重,最终加权得到质量评分S1。

[0128] 其中,归一化权重的具体获得方式,可以参考如下实例操作:

[0129] 当1~5级mAP评分(同程度分)分别为0.2、0.4、0.6、0.8、1时,评价算法效果相当;权重

$w_n = \left( \frac{1}{\text{同程度分 mAP}_n} \right)$ ,因此,对同程度分(0.2、0.4、0.6、0.8、1)取倒数,就推算出权重为(5,2.5,1.66666667,1.25,1)。

[0130] 当1-5级mAP评分分别为1、1、1、1、1时,得到质量总评分应为1,而不是5。因此,需要归一化权重,方法: $w \text{ 和} = (5 + 2.5 + 1.66666667 + 1.25 + 1) = 11.41666667$ ,由

此,权重 $w = (5, 2.5, 1.66666667, 1.25, 1) / 11.41666667 = (0.4379562, 0.2189781, 0.1459854, 0.10948905, 0.08759124)$ 。

[0131] 最后,对各等级的mAP\*w取加和,得最终评分S,公式,其中,n为5,分别对应不同等级。

[0132] 试验:当mAP为(0.2, 0.4, 0.6, 0.8, 1)时,各自评分应为0.08759124(程度分相同),加和为0.4379562。

[0133] 结果:代入权重归一化公式得到总评分为0.4379562,正确。0.4379562为一般效果分,若想要更高分,则得加强对强噪声图像的训练。

[0134] 试验2:当mAP为(0.7, 0.7, 0.7, 0.7, 0.7)时,总评分应为0.7;

[0135] 结果:代入权重归一化公式得到总评分为0.7,正确。

[0136] (2) 总体图像分辨率评分

[0137] 在TG测试集中按分辨率大、中、小分为3类,再分别对此3类进行评分计算mAP,并最终平均加权得到分辨率评分S2。

[0138] (3) 总体目标质量评分

[0139] 选择TG试集,在总体测试结果中选择此测试集的结果。按目标质量好与差,分为2个集合,再分别对此2类进行评分计算mAP,并最终通过归一化权重进行加权得到分辨率评分S3。

[0140] 这里的归一化权重的获得方式,如前述的总体图像质量评分中的操作方式。设置目标质量好的mAP同程度分为1分,质量差的mAP同程度分为0.5分,则权重为(1/1,1/0.5),再进行归一化,相除:(1/1,1/0.5)/(1/1+1/0.5),即权重分别为(0.33333,0.6666666)。

[0141] (4) 总体目标大小评分

[0142] 选择TG试集,在总体测试结果中选择此测试集的结果。按目标大、中、小,分为3个集合,再分别对此3类进行评分计算mAP,并最终平均加权得到分辨率评分S4。

[0143] (5) 总体综合评分

[0144] 按照以上图像不同性质对算法进行评分得到S1~S4,最终对其平均加权得到总分S总。

[0145] 二、若选择需要进行应用场景评分,除应用场景图像质量评分采用等级最高的两个数据集中的一个数据集进行测评,其他评分只用对专家/用户参数中设置的应用环境参

数下的测试集进行评分。例如,系统跟原测试集的图集属性情况,安装分辨率大中小或者多余三级的情况进行分辨率分类,同时根据目前质量好与差分为至少2级以上,根据目标大小分为多级。用户根据需要测试的应用场景情况,选择对应的分辨率、目标质量、目标大小等级,以对应出需要的分辨率范围、目标质量范围、目标大小调。

[0146] 例如用户将参数设置为:分辨率等级2,目标质量等级2,目标大小等级2,就是选择了这些等级对应的分辨率范围、目标质量范围、目标大小范围。分别计算以上参数下的评分,再加上质量评分,输入为具有专家参数性质的图片,输出为各情况下的评分,最终加权得到总评分。

[0147] (1)场景图像质量评分

[0148] 对TG测试集进行加噪处理,分别提取5种质量等级的5个测试集T1~T5。

[0149] 若各测试集的质量等级排序为:T1<T2<T3<T4<T5。分别采用T4或T5进行测试,计算T4与P4的mAP4,或,计算T5与P5的mAP5,得出的评分,作为场景图像质量评分S1。

[0150] (2)场景图像分辨率评分

[0151] 用户根据自己的需求,输入分辨率为中的参数,在具体实例中,设分辨率等级2,根据分辨率等级2对应生成属于分辨率等级2的场景图像分辨率测试集,并计算mAP评分,获得场景图像分辨率评分S2。

[0152] (3)场景目标质量评分

[0153] 只对根据用户输入的目标质量参数,例如设为目标质量等级2,则生成目标质量为2对应的目标质量范围的测试集,并进行测试,计算mAP评分,获得场景目标质量评分S3。

[0154] (4)场景目标大小评分

[0155] 只对根据用户输入的目标大小参数,例如设目标大小等级为2,则生成目标大小等级为2对应的目标大小范围的测试集,进行测试,计算mAP评分,获得场景目标大小评分S4。

[0156] (5)应用场景评分

[0157] 按照以上图像不同性质对算法进行评分得到S1~S4,最终对其平均加权得到总分S总。

[0158] 三、某些算法想对某些特定场景项目进行单独测试,比如只对算法进行目标大小的评分,或分辨率大小评分,又或以上两项或其他项目进行评分。使用者可按照自己的想法,进行自由评分。系统根据用户输入的参数进行对应自由测试场景/测试集的生成。比如输入为具有小目标的图片,输出为小目标评分。

[0159] (1)数据准备

[0160] 用户在自由场景预设单元中输入自己想要的评分场景,具体通过输入对应的参数,由自由场景预设单元对应生成自由数据集。

[0161] (2)自由评分

[0162] 对测试集进行测试,并进行自由评分,得到一项或多项评分结果,以供算法模型的优化。若为多项,则采用平均加强获得总体自由评分。

[0163] 以上述实例及场景参数作为评价系统/方法,采用yolov3\_tiny版对kitti数据集进行训练,并对测试集进行目标检测对比如下(其中,实例中,应用场景评分,采用T4等级计算应用场景图像质量评分,采用分辨率等级2、目标质量等级2、目标大小登记2的参数计算其他评分):

算法		评分	总分
传统算法 从预测标签中提出未做质量算法处理的数据的标签（质量等级为5的图片），与真实标签计算 mAP。		0.723	0.723
[0164] 总体综合 评分	总体图像质量评分	0.1588	0.4674
	总体图像分辨率评分	0.625	
	总体目标质量评分	0.5354	
	总体目标大小评分	0.550820	
应用 场景 评分	场景图像质量评分 (质量等级4)	0.5254	0.720825
	场景图像分辨率评分 (分辨率等级为2)	0.731	
	场景目标质量评分 (目标质量等级为2)	0.8259	
	场景目标大小评分 (目标大小等级为2)	0.801	
自由评分	自由图像分辨率评分	得到图像分辨率等级为1的分数: 0.684。	

[0165] 通过本系统可以同时进行总体综合评分、应用场景评分、自由评分。如上表所列的实例结果,则可以:

[0166] 通过总体综合评分与传统评分对比:可发现此待测算法对无噪声图像上表现很好,但对加噪图像目标检测敏感,对大分辨率敏感,对小目标敏感,对遮挡敏感,应加强以上方面属性的图像的训练;

[0167] 通过应用场景评分与传统评分对比:可发现此待测算法在特定应用场景(即预设分辨率等级、目标质量等级、目标大小等级各自对应的测试集/测试图集)上表现不错,还需要加强对应用场景的噪声图像的训练;

[0168] 自由评分属于特定的测评,不参与比对。

[0169] 通过本发明提供的系统/方法,能给出总体综合评分、应用场景评分、综合自由评分,不仅能客观反映出算法的好坏,还能反映出算法的好坏领域,而且能给算法设计者明确的优化方向,使算法优化更有针对性和目的性。

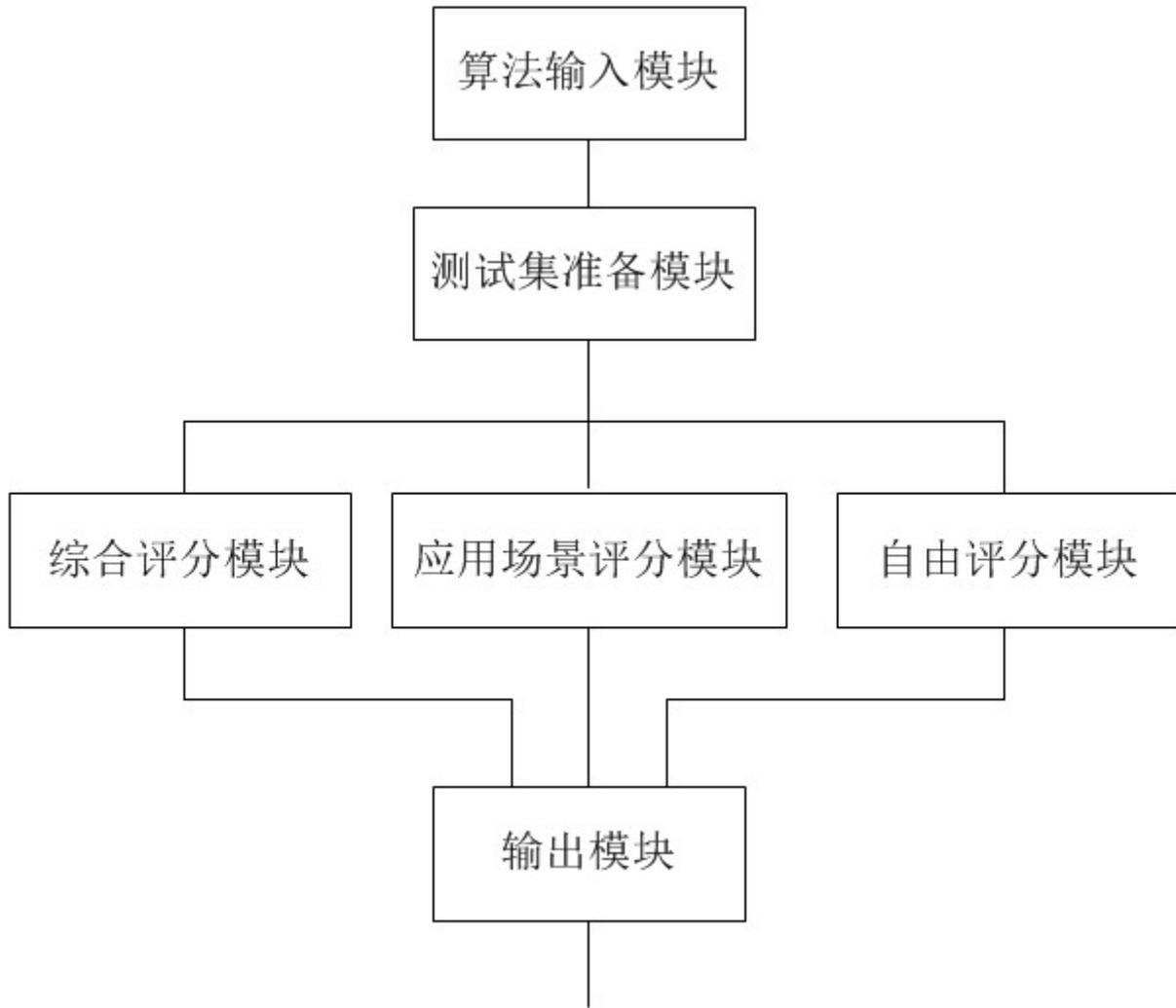


图1

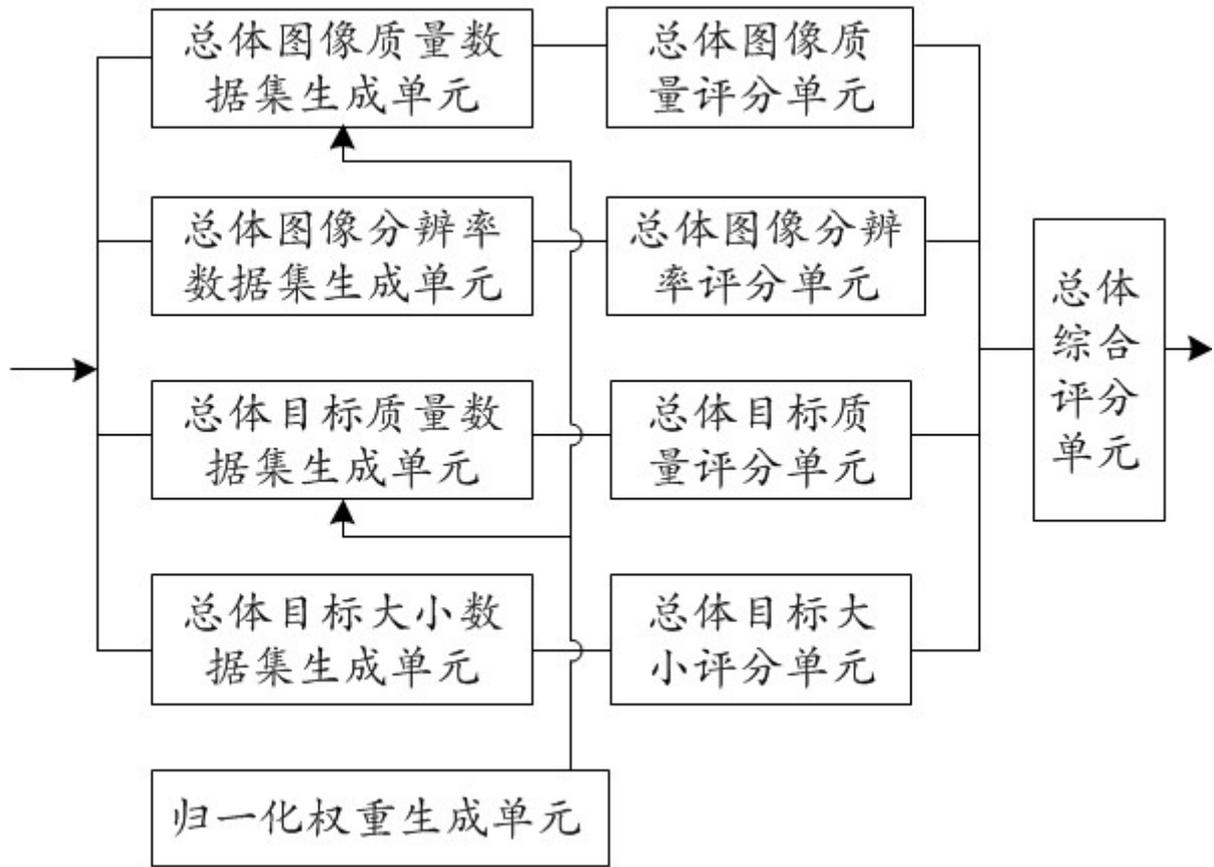


图2

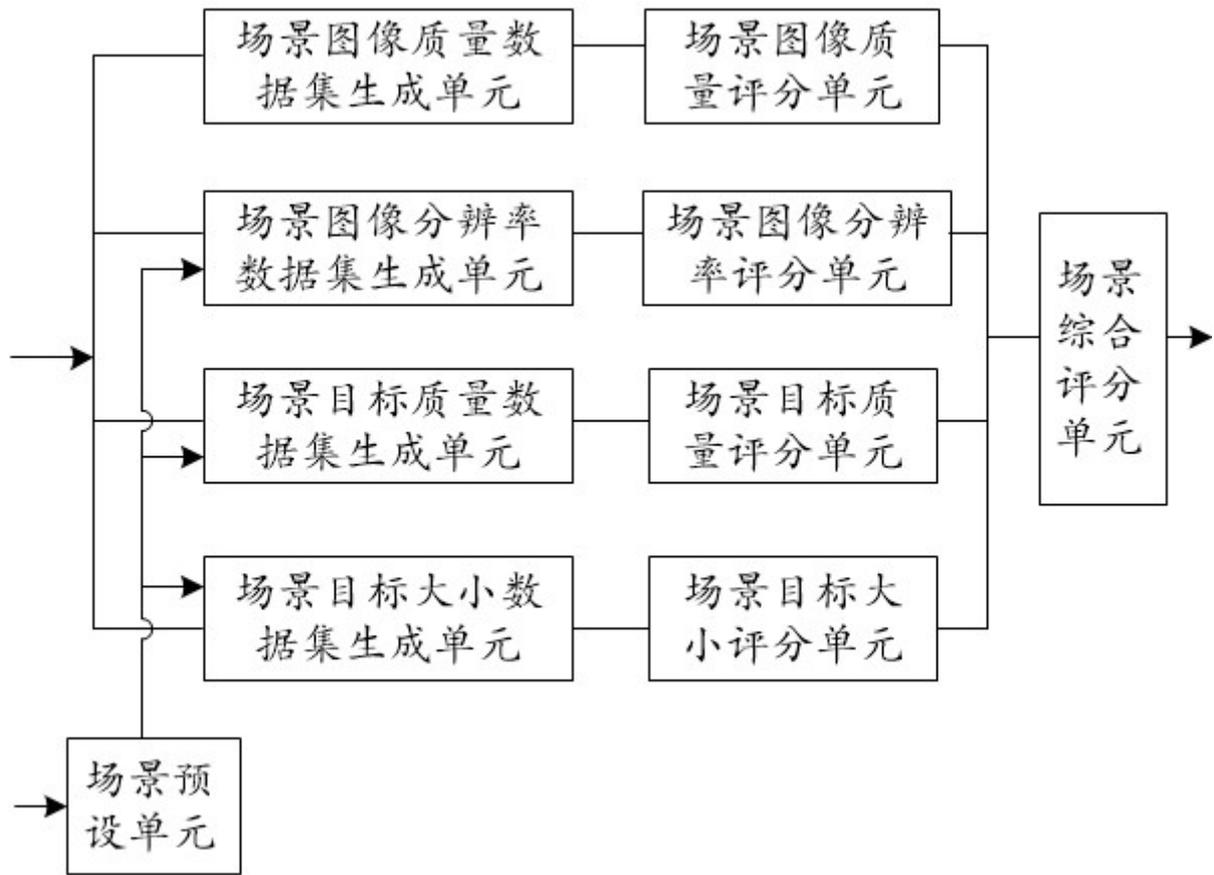


图3

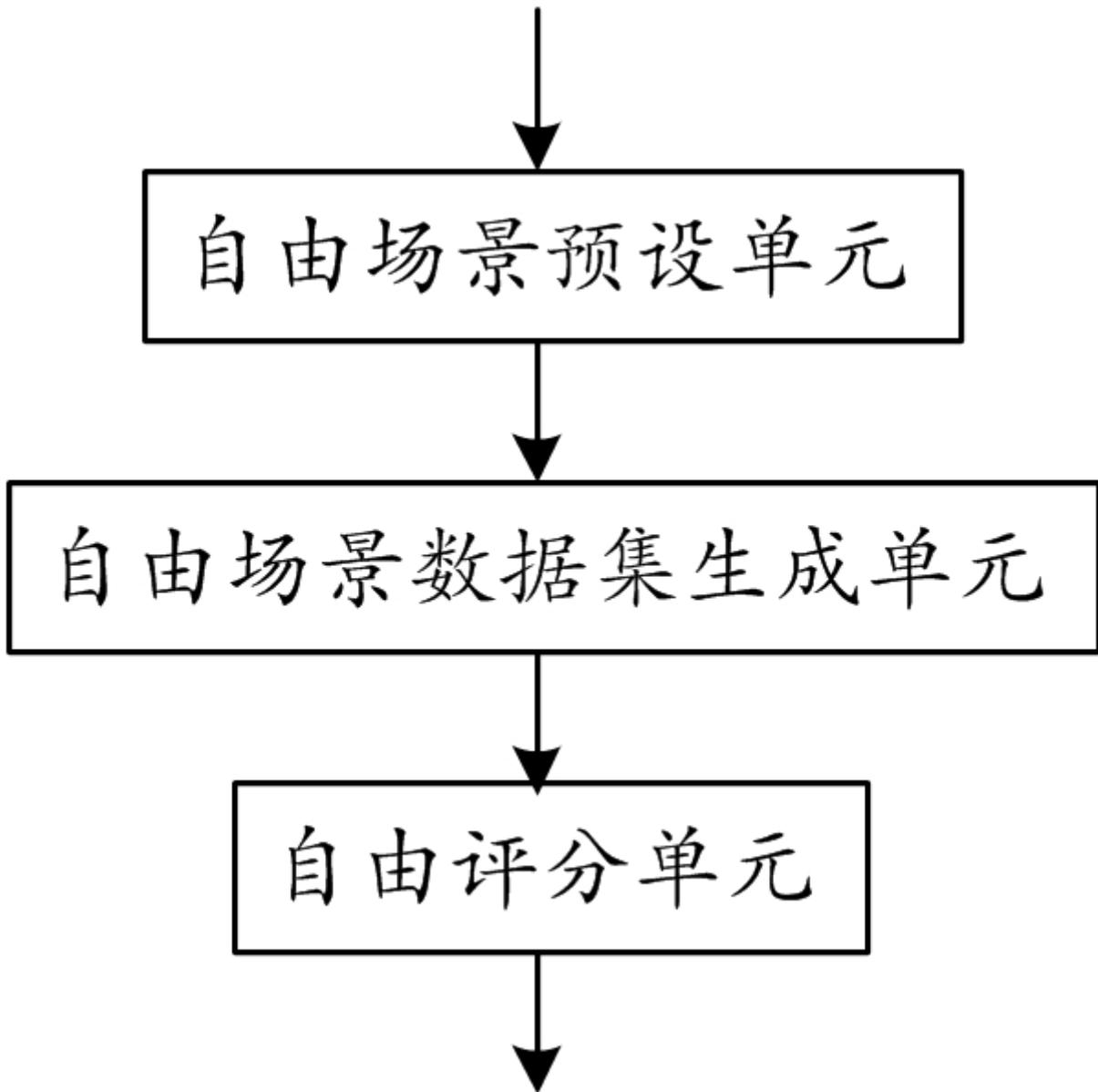
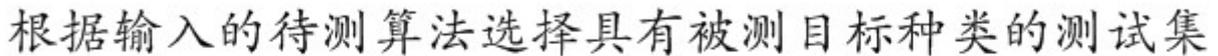
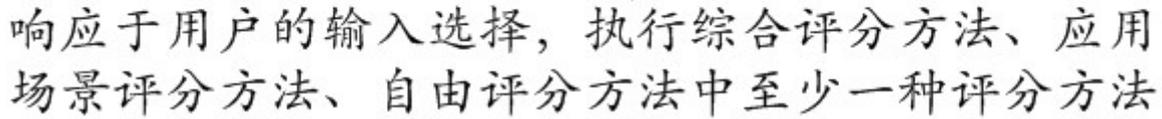


图4

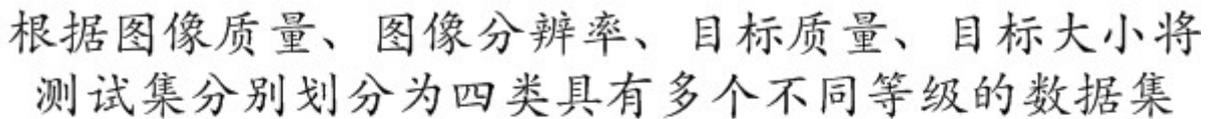


根据输入的待测算法选择具有被测目标种类的测试集

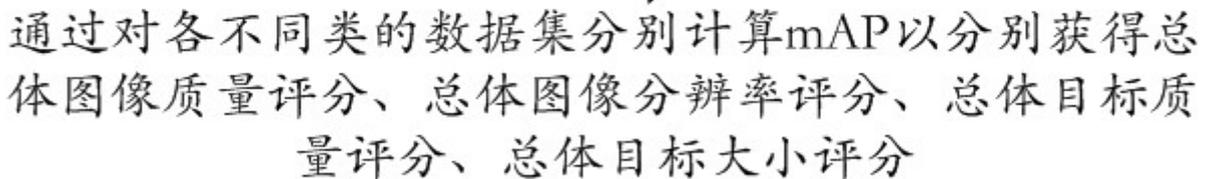


响应于用户的输入选择，执行综合评分方法、应用场景评分方法、自由评分方法中至少一种评分方法

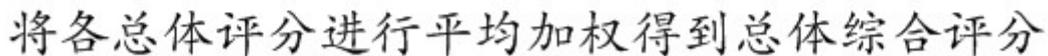
图5



根据图像质量、图像分辨率、目标质量、目标大小将测试集分别划分为四类具有多个不同等级的数据集



通过对各不同类的数据集分别计算mAP以分别获得总体图像质量评分、总体图像分辨率评分、总体目标质量评分、总体目标大小评分



将各总体评分进行平均加权得到总体综合评分

图6

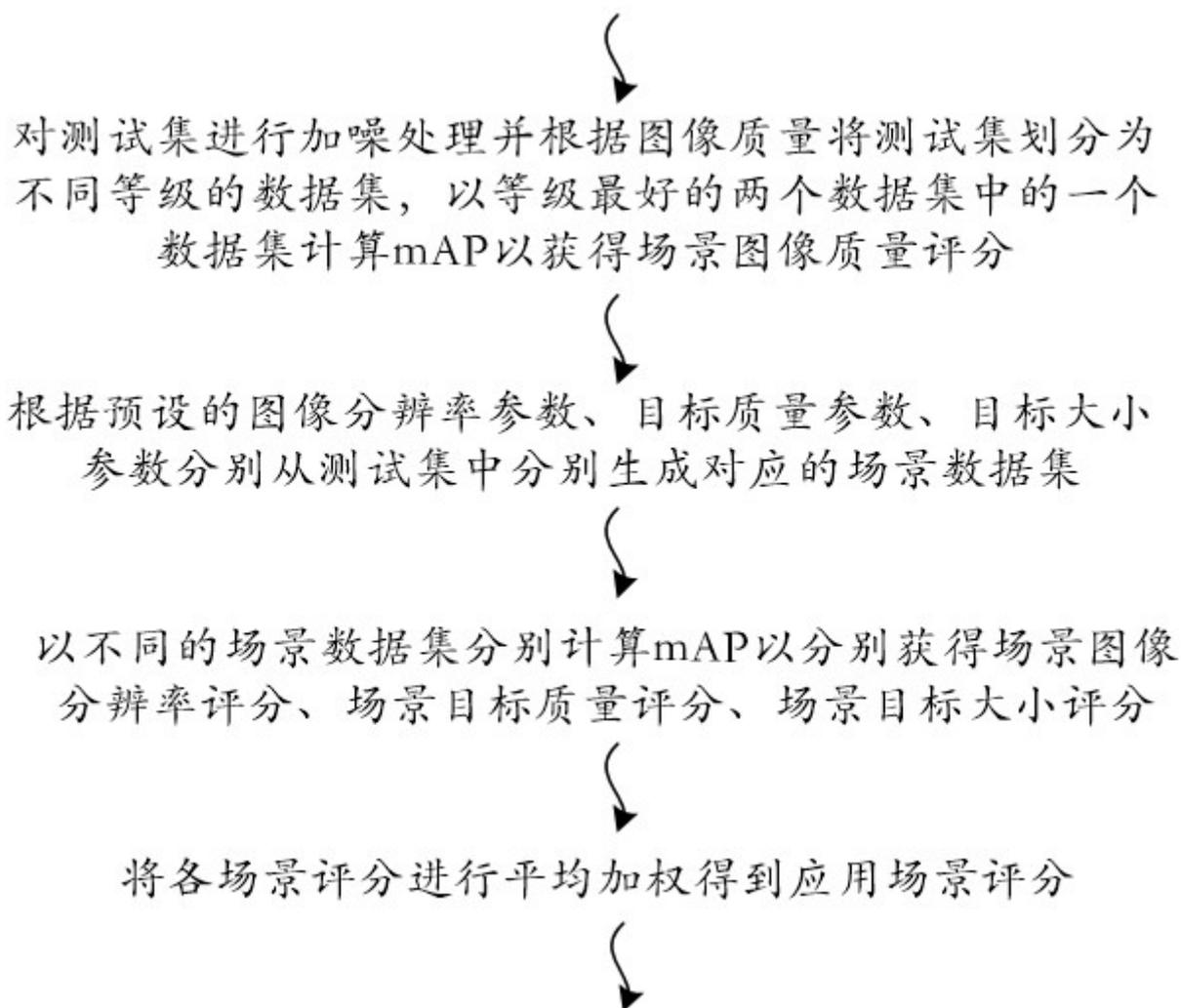


图7

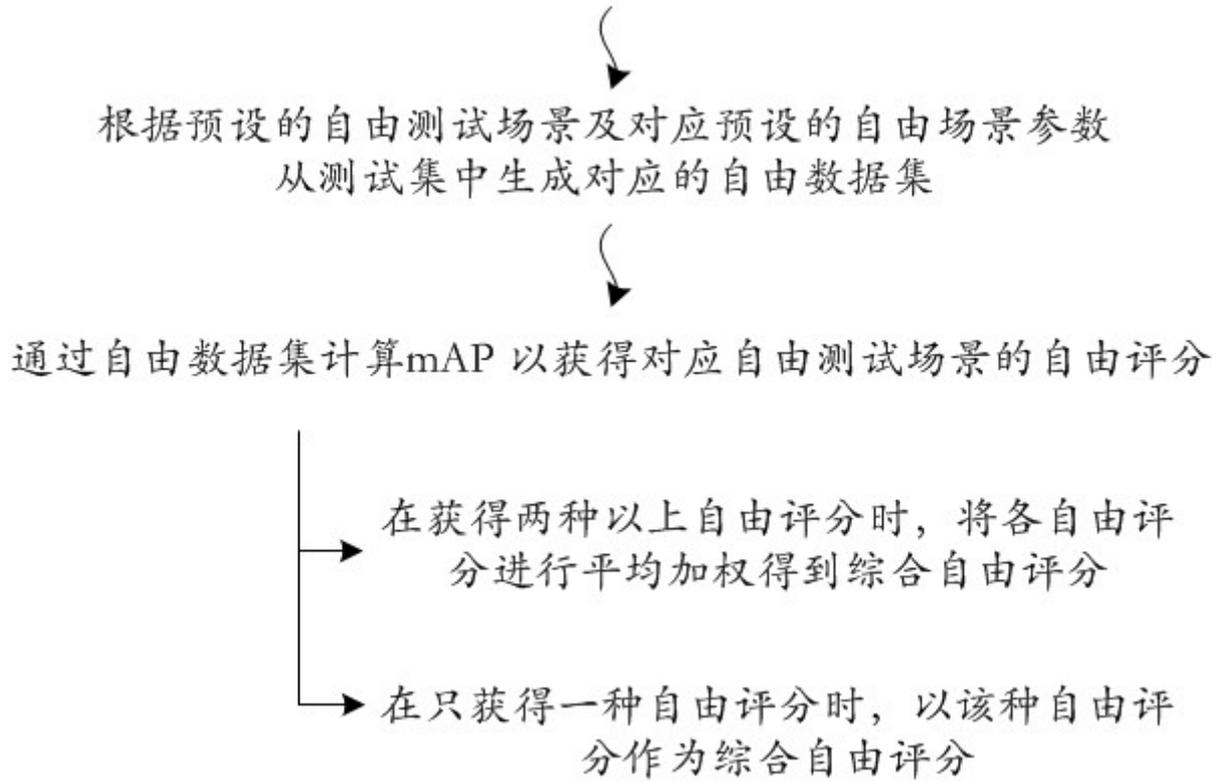


图8