



(12) 发明专利申请

(10) 申请公布号 CN 114463759 A

(43) 申请公布日 2022. 05. 10

(21) 申请号 202210386933.3

G06V 10/80 (2022.01)

(22) 申请日 2022.04.14

G06V 10/82 (2022.01)

(71) 申请人 浙江霖研精密科技有限公司

地址 313000 浙江省湖州市南浔区练市镇
茹家甸路100号

(72) 发明人 张晓武 陈斌 李伟 徐朝彬

(74) 专利代理机构 成都君合集专利代理事务所
(普通合伙) 51228

专利代理师 尹新路

(51) Int. Cl.

G06V 30/148 (2022.01)

G06N 3/08 (2006.01)

G06N 3/04 (2006.01)

G06K 9/62 (2022.01)

G06V 10/774 (2022.01)

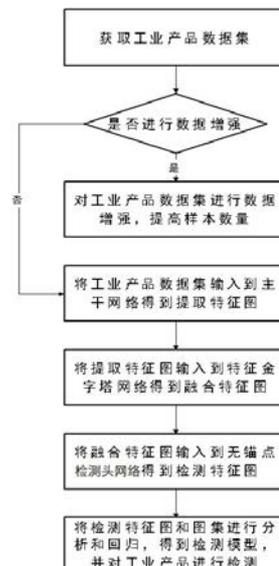
权利要求书2页 说明书9页 附图4页

(54) 发明名称

一种基于无锚框算法的轻量化文字检测方法
及装置

(57) 摘要

本发明涉及无锚框算法技术领域,公开了一种基于无锚框算法的轻量化文字检测方法,包括:步骤S1.获取工业产品上文字被框选的工业产品图像样本作为工业产品数据集,对工业产品数据集进行标注并将工业产品数据集分为测试集和训练集;步骤S2.将无锚点检测头网络工业产品数据集的训练集输入网络模型获取训练后的网络模型,无锚点检测头网络网络模型包括依次连接的主干网络、特征金字塔网络和无锚点检测头网络;步骤S3.使用无锚点检测头网络测试集测试训练后的网络模型得到文字检测模型;步骤S4.将待处理数据输入文字检测模型进行检测并输出检测结果。本发明还公开了一种基于无锚框算法的轻量化文字检测装置。



1. 一种基于无锚框算法的轻量化文字检测方法,其特征在于,包括以下步骤: 步骤S1: 获取工业产品上文字被框选的工业产品图像样本作为工业产品数据集,对工业产品数据集进行标注,并将工业产品数据集分为测试集和训练集; 步骤S2: 基于Anchor-Free算法将所述工业产品数据集的训练集输入网络模型获取训练后的网络模型,所述网络模型包括依次连接的主干网络、特征金字塔网络和无锚点检测头网络,所述无锚点检测头网络引入了广义焦点损失,并去掉了FCOS检测器使用的焦点损失中的Centerness分支; 步骤S3: 使用所述测试集测试训练后的网络模型得到文字检测模型; 步骤S4: 将待处理数据输入文字检测模型进行检测并输出检测结果。

2. 根据权利要求1所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S1包括:

对所述工业产品图像样本进行数据增强处理得到增强后的工业产品图像样本作为工业产品数据集;

所述数据增强处理的方法包括依次对工业产品图像样本进行镜像操作和尺寸调整操作,所述尺寸调整操作为随机扩充或者随机裁剪。

3. 根据权利要求1所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S2包括: 步骤S21: 将所述工业产品数据集的训练集输入到主干网络获取不同尺寸的提取特征图; 步骤S22: 将所述提取特征图输入到特征金字塔网络进行多级预测获取融合特征图; 步骤S23: 将所述融合特征图输入到无锚点检测头网络进行分析和回归计算获取检测特征图; 步骤S24: 计算无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差,根据所述误差采用反向传播算法对所述网络模型进行训练,把检测特征图的每个位置作为训练样本,对于检测特征图中的每个位置对应原图的边框都进行回归,最终获取训练后的网络模型。

4. 根据权利要求3所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S21中的主干网络包括ShuffleNetV2网络;

所述ShuffleNetV2网络从前至后依次设置卷积层、批标准化层、激活函数层和若干个ShuffleNet基本单元,并去掉了最后一层卷积,抽取8、16、32倍下采样的特征对工业产品数据集的训练集进行多尺度的特征融合,获取不同尺寸的提取特征图;

所述ShuffleNet基本单元包括从前至后依次设置的分组卷积层、通道随机混合操作层、深度可分离卷积层、卷积层和特征相加层。

5. 根据权利要求3所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S22中的特征金字塔网络包括PAN网络;

所述PAN网络中去掉了所有卷积层,保留从主干网络中特征提取后的1x1卷积来进行特征通道维度的对齐;

所述PAN网络中的上采样和下采样均使用插值完成,并且将多尺寸的提取特征图直接相加。

6. 根据权利要求3所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S23中的无锚点检测头网络包括:

所述无锚点检测头网络采用了共享权重的检测头,对从特征金字塔网络出来的多尺度特征图使用同一组卷积预测检测框,然后每一层使用一个可学习的缩放值作为系数,对预

测出来的框进行缩放,并且在检测头上使用了组归一化层的处理方式。

7. 根据权利要求3所述的一种基于无锚框算法的轻量化文字检测方法,其特征在于,所述步骤S24中计算无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差的方法包括:

无锚点检测头网络的损失值GFL代表了无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差,无锚点检测头网络的损失值GFL表示为:

$$GFL(p_{y_l}, p_{y_r}) = -|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta ((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}));$$

其中, y_l 和 y_r 表示预测模型的两个变量,这两个变量的概率值为 p_{y_l} 和 p_{y_r} , $p_{y_l} + p_{y_r} = 1$, 连续标签 y 满足 $y_l < y < y_r$, 最终预测为两个变量和概率的线性组合: $\hat{y} = y_l p_{y_l} + y_r p_{y_r}$;

$|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta$ ($\beta \geq 0$) 是调节因子;

$((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}))$ 是交叉熵的完全表现形式。

8. 一种基于无锚框算法的轻量化文字检测装置,其特征在于,包括采集模块、训练模块和检测模块:

所述采集模块用于采集工业产品上文字被框选的工业产品图像样本作为工业产品数据集,并将工业产品数据集分为测试集和训练集;

所述训练模块用于采集训练集训练网络模型,所述网络模型包括主干网络、特征金字塔网络和无锚点检测网络;

所述检测模块用于使用所述测试集测试训练后的网络模型得到文字检测模型,所述文字检测模型用于进行文字检测。

9. 根据权利要求8所述的一种基于无锚框算法的轻量化文字检测装置,其特征在于,包括: 所述训练模块将训练集输入进主干网络中进行特征提取得到不同尺寸的特征图,再采用特征金字塔网络对不同尺寸的特征图进行多级预测,最后在无锚点检测网络中对不同尺寸的特征图进行分类和回归处理,把特征图的每个位置作为训练样本,对于特征图中的每个位置对应原图的边框都进行回归,最终得到训练好的网络模型。

10. 根据权利要求8所述的一种基于无锚框算法的轻量化文字检测装置,其特征在于,所述训练模块包括:

所述训练模块采用了训练辅助模块,并结合了动态的标签匹配策略;

所述训练模块包括从前至后依次设置的卷积层、组归一化层,并在不同尺度的特征图间共享参数;

所述分标签匹配策略采用自适应正负anchor选择的匹配策略,为每一层特征图动态选取匹配样本。

一种基于无锚框算法的轻量化文字检测方法及装置

技术领域

[0001] 本发明涉及无锚框算法技术领域,具体地说,是一种基于无锚框算法的轻量化文字检测方法及装置,用于对工业产品的文字进行实时检测。

背景技术

[0002] 工业产品检测方法主要分为传统方法和人工智能方法。传统方法也分两种,一种完全由人眼检测,这种方法检测结果不稳定,检测人员的主观因素占比很大,而且随着产品产量的增加,检测人员会出现视觉疲劳,造成大量误检;另外一种是对工业产品提取手工特征进行分类,是一种应用传统图像处理衍生的技术,但是一些产品上面的文字或图案,由于模具不一致以及图像模糊等原因,使用传统的图像匹配算法容易出错,造成误检。人工智能方法是指使用基于深度学习的方法进行检测,通过深度神经网络模型对采集的工业产品图像数据进行分析,定位出检测目标,在数据量较大且复杂的情况下,也能通过加深加宽网络模型,提高模型的特征表达能力,从而精准检测产品,得到了令人满意的检测效果。基于深度学习的检测技术主要分为两阶段算法和单阶段算法,两阶段算法中代表算法有Fater R-CNN、MTCNN等,这类基于Anchor-Based算法主要将检测过程分成生成候选区域和基于候选区域检测分类,具有高精度的优势。而单阶段代表算法有YOLOV3、SSD等,这类算法基于Anchor-Free,没有候选区域部分的计算,直接在特征图上定位分类,具有较高的实时性。

[0003] 目前,大多数基于深度学习的工业产品检测,为了能够得到精度较高的模型,模型结构比较复杂,从而产生的模型比较笨重,这直接影响了在实际工业应用中由于设备的限制,不能够有效的实施。因此,亟需一种部署简单的轻量化工业产品检测方案,使得模型能够在移动端部署,以达到实时对工业产品上文字位置准确检测的性能。

发明内容

[0004] 本发明的目的在于提供一种基于无锚框算法的轻量化文字检测方法及装置,具有实时对工业产品上文字位置准确检测的效果。

[0005] 本发明通过下述技术方案实现:一种基于无锚框算法的轻量化文字检测方法,包括以下步骤:

步骤S1:获取工业产品上文字被框选的工业产品图像样本作为工业产品数据集,对工业产品数据集进行标注,并将工业产品数据集分为测试集和训练集;

步骤S2:基于Anchor-Free算法将所述工业产品数据集的训练集输入网络模型获取训练后的网络模型,所述网络模型包括依次连接的主干网络、特征金字塔网络和无锚点检测头网络,所述无锚点检测头网络引入了广义焦点损失,并去掉了FCOS检测器使用的焦点损失中的Centerness分支;

步骤S3:使用所述测试集测试训练后的网络模型得到文字检测模型;

步骤S4:将待处理数据输入文字检测模型进行检测并输出检测结果。

[0006] 本发明旨在在工业场景中,提供一种能够对工业产品上文字位置准确检测的轻量

化模型,模型能够在移动端达到实时检测的性能,如在实际工业应用中由于设备的限制,深度学习模型不能够有效实施。本发明中的Anchor-Free算法,即无锚框算法,是由于其结构简单,运算效率高,且对移动端检测更加友好,并引入了广义焦点损失,其去掉了FCOS检测器使用的焦点损失中的Centerness分支,减少了检测器的计算开销,加快检测速度。

[0007] 为了更好地实现本发明,进一步地,步骤S1包括:

对所述工业产品图像样本进行数据增强处理得到增强后的工业产品图像样本作为工业产品数据集;

所述数据增强处理的方法包括依次对工业产品图像样本进行镜像操作和尺寸调整操作,所述尺寸调整操作为随机扩充或者随机裁剪。

[0008] 为了更好地实现本发明,进一步地,步骤S2包括:

步骤S21:将所述工业产品数据集的训练集输入到主干网络获取不同尺寸的提取特征图;

步骤S22:将所述提取特征图输入到特征金字塔网络进行多级预测获取融合特征图;

步骤S23:将所述融合特征图输入到无锚点检测头网络进行分析和回归计算获取检测特征图;

步骤S24:计算无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差,根据所述误差采用反向传播算法对所述网络模型进行训练,把检测特征图的每个位置作为训练样本,对于检测特征图中的每个位置对应原图的边框都进行回归,最终获取训练后的网络模型。

[0009] 为了更好地实现本发明,进一步地,步骤S21中的主干网络包括ShuffleNetV2网络;

所述ShuffleNetV2网络从前至后依次设置卷积层、批标准化层、激活函数层和若干个ShuffleNet基本单元,并去掉了最后一层卷积,抽取8、16、32倍下采样的特征对工业产品数据集的训练集进行多尺度的特征融合,获取不同尺寸的提取特征图;

所述ShuffleNet基本单元包括从前至后依次设置的分组卷积层、通道随机混合操作层、深度可分离卷积层、卷积层和特征相加层。

[0010] 为了更好地实现本发明,进一步地,步骤S22中的金字塔网络包括PAN网络;

所述PAN网络中去掉了所有卷积层,保留从主干网络中特征提取后的1x1卷积来进行特征通道维度的对齐;

所述PAN网络中的上采样和下采样均使用插值完成,并且将多尺寸的提取特征图直接相加。

[0011] 为了更好地实现本发明,进一步地,步骤S23中的无锚点检测头网络包括:

所述无锚点检测头网络采用了共享权重的检测头,对从特征金字塔网络出来的多尺度特征图使用同一组卷积预测检测框,然后每一层使用一个可学习的缩放值作为系数,对预测出来的框进行缩放,并且在检测头上使用了组归一化层的处理方式。

[0012] 为了更好地实现本发明,进一步地,步骤S24中计算无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差的方法包括:无锚点检测头网络的损失值GFL代表了无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差,无锚点检测

头网络的损失值GFL表示为：

$$GFL(p_{y_l}, p_{y_r}) = -|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta ((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}));$$

其中, y_l 和 y_r 表示预测模型的两个变量,这两个变量的概率值为 p_{y_l} 和 p_{y_r} , $p_{y_l} + p_{y_r} = 1$,连续标签 y 满足 $y_l < y < y_r$,最终预测为两个变量和概率的线性组合: $\hat{y} = y_l p_{y_l} + y_r p_{y_r}$;

$|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta$ ($\beta \geq 0$)是调节因子;

$((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}))$ 是交叉熵的完全表现形式。

[0013] 为了更好地实现本发明,进一步地,本发明还提供了一种基于无锚框算法的轻量化文字检测装置,包括采集模块、训练模块和检测模块:

所述采集模块用于采集工业产品上文字被框选的工业产品图像样本作为工业产品数据集,并将工业产品数据集分为测试集和训练集;

所述训练模块用于采集训练集训练网络模型,所述网络模型包括主干网络、特征金字塔网络和无锚点检测网络;

所述检测模块用于使用所述测试集测试训练后的网络模型得到文字检测模型,所述文字检测模型用于进行文字检测。

[0014] 为了更好地实现本发明,进一步地,训练模块将训练集输入进主干网络中进行特征提取得到不同尺寸的特征图,再采用特征金字塔网络对不同尺寸的特征图进行多级预测,最后在无锚点检测网络中对不同尺寸的特征图进行分类和回归处理,把特征图的每个位置作为训练样本,对于特征图中的每个位置对应原图的边框都进行回归,最终得到训练好的网络模型。

[0015] 为了更好地实现本发明,进一步地,所述训练模块包括:

所述训练模块采用了训练辅助模块,并结合了动态的标签匹配策略;

所述训练模块包括从前至后依次设置的卷积层、组归一化层,并在不同尺度的特征图间共享参数;

所述分标签匹配策略采用自适应正负anchor选择的匹配策略,为每一层特征图动态选取匹配样本。

[0016] 本发明与现有技术相比,具有以下优点及有益效果:

(1) 本发明解决了在解决工业产品在生产过程中,由于产品表面会受到药粉等不可控的脏污,从而导致对产品检测出现误检的问题,并使用Anchor-Free算法,即无锚框算法,使得模型达到轻量级;

(2) 本发明通过引入特征融合模块PAN和无锚点检测头网络,结合主干网络ShuffleNetV2,构成实时轻量化无锚点的文字检测网络,提高了检测网络的运行效率,减少了模型参数,加快检测速度;

(3) 本发明能够在嵌入式平台和GPU计算平台等多平台实现实时目标检测。通过验证和评估,检测精度可以达到99.6%的mAP@0.5,检测速度在GTX 3060上每张1400*1400分辨率的图片不超过20ms。

附图说明

[0017] 本发明结合下面附图和实施例做进一步说明,本发明所有构思创新应视为所公开内容和本发明保护范围。

[0018] 图1为本发明实施例提供的一种基于无锚框算法的轻量化文字检测装置的流程图。

[0019] 图2为本发明实施例提供的基于无锚框算法的轻量化文字检测装置的网络结构图。

[0020] 图3为为本发明实施例提供的ShuffleNetV2网络中ShuffleNet结构的示意图。

[0021] 图4为本发明实施例提供的基于无锚框算法的轻量化文字检测装置的结构框图。

具体实施方式

[0022] 为了更清楚地说明本发明实施例的技术方案,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,应当理解,所描述的实施例仅仅是本发明的一部分实施例,而不是全部的实施例,因此不应被看作是对保护范围的限定。基于本发明中的实施例,本领域普通技术工作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。

[0023] 在本发明的描述中,需要说明的是,除非另有明确的规定和限定,术语“设置”、“相连”、“连接”应做广义理解,例如,可以是固定连接,也可以是可拆卸连接,或一体地连接;可以是机械连接,也可以是电连接;也可以是直接相连,也可以是通过中间媒介间接相连,可以是两个元件内部的连通。对于本领域的普通技术人员而言,可以根据具体情况理解上述术语在本发明中的具体含义。

[0024] 实施例1:

本实施例的一种基于无锚框算法的轻量化文字检测方法,如图1-图3所示,目前大多数优秀检测算法都是Anchor-Based,也就是基于锚点的检测算法,但由于锚点先验大都需要通过人为进行设计,导致基于锚点的检测算法存在诸多缺点,例如计算冗余、参数量大,检测速度慢等,随着基于Anchor-Free,也就是无锚框算法的检测精度快速提高,基于无锚点的检测算法被越来越多的学者所关注,产生了很多优秀算法,此外更有研究表明,目标检测数据集中正负样本的定义是两者性能优劣的根本区别,因此设计一个合理快速的基于Anchor-Free的轻量化文字检测装置是具有一定的可行性的,综上,本实施例提供了一种基于Anchor-Free的轻量化文字检测方法,如图1所示,包括:获取工业产品数据集;工业产品数据集为产品上文字被框选的工业产品图像。将工业产品数据集输入网络模型得到多层特征图;网络模型包括依次连接的主干网络ShuffleNetV2、特征金字塔网络和无锚点检测头网络;主干网络ShuffleNetV2的归一化方法包括组归一化和批标准化,组归一化和批标准化按照权重结合,权重可通过训练进行优化。对于多层特征图中的每一层都连接上一个无锚点检测头进行分析和回归,计算特征图与图集中各图之间的误差,根据误差采用反向传播算法对网络模型进行训练得到文字检测模型,文字检测模型用于进行文字检测。

[0025] 本发明通过获取训练图像数据集,可以对数据集进行数据增强,以特征提取网络ShuffleNetV2为主干网络,并结合特征金字塔网络和无锚点检测头网络,构建检测模型;通过训练图像数据集对检测模型进行训练,得到训练后的检测模型;将含有待检测目标的图

像作为输入,通过训练好的检测模型对图像中的待检测目标进行检测,本发明旨在解决工业产品在生产过程中,由于产品表面会受到药粉等不可控的脏污,从而导致对产品检测出现误检的问题,因此获取工业产品上文字被框选的工业产品图像作为工业产品数据集。

[0026] 实施例2:

本实施例在实施例1的基础上做进一步优化,在本实施例中,获取工业产品数据集的过程中进行了数据增强,对工业产品数据集进行数据增强的过程其实是对工业产品图像样本进行数据增强的过程,是为了提高样本数量,之后再继续进行后续操作。

[0027] 在本实施例中标注的信息包含检测文字的类别和位置信息,标注的信息比例以自定义的比例将工业产品数据集分为测试集和训练集,经常设定为8:2,训练集和测试集包含的标注信息是一样的。因为划分数据集后训练集数量足够多,足够用于训练,所以将训练集输入网络模型进行训练。

[0028] 本实施例的其他部分与实施例1相同,故不再赘述。

[0029] 实施例3:

本实施例在上述实施例1或2的基础上做进一步优化,本实施例在实现过程中,首先获取工业产品训练图像数据集,接着以主干网络进行特征提取得到特征图,然后根据特征图结合特征金字塔网络和无锚点检测头网络进行多层预测,构建检测模型;通过训练图像数据集对检测模型进行训练,得到训练后的检测模型;将含有待检测目标的图像作为输入,通过训练好的检测模型对图像中的待检测目标进行检测。

[0030] 实施例的一种基于无锚框算法的轻量化文字检测方法,包括如下阶段和步骤:

阶段1:对工业产品数据集进行数据增强处理得到增强后的工业产品数据集,具体包括:依次对工业产品数据集进行镜像操作和尺寸调整操作,尺寸调整操作为随机扩充或者随机裁剪;

阶段2:构建基于Anchor-Free的轻量化文字检测模型,其包括以下3个步骤:

首先,搭建主干网络ShuffleNetV2,ShuffleNetV2网络包括从前至后依次设置的卷积层、批标准化层、激活函数层以及若干个ShuffleNet基本单元,并且去掉了最后一层卷积,并且抽取8、16、32倍下采样的特征进行多尺度的特征融合。

[0031] 其次,搭建特征金字塔网络,采用了特征融合模块PAN结构,去掉了PAN中的所有卷积层,保留从主干网络中特征提取后的1x1卷积来进行特征通道维度的对齐,其中,上采样和下采样均使用插值完成,并且将多尺寸的特征图直接相加,使得整个模块的计算量非常小。如图2所示,对P5通过上采样和下采样操作后分别得到了P3、P4和P6、P7的不同层次的特征图,然后将通过主干网络产生的特征层C3、C4、C5与对应的P3、P4、P5进行特征融合。

[0032] 最后,搭建检测头结构,无锚点检测头网络采用了共享权重的检测头,采用了FCOS系列中的检测头,使用了共享权重的检测头,即对FPN出来的多尺度Feature Map使用同一组卷积预测检测框,然后每一层使用一个可学习的Scale值作为系数,对预测出来的框进行缩放,这能够将检测头的参数量降低为不共享权重状态下的1/5,但是由于移动端模型推理由CPU进行计算,共享权重并不会对推理过程进行加速,而且在检测头非常轻量,共享权重使得其检测能力进一步下降,因此本发明可以选择对每一层特征使用一层卷积。同时,检测头使用了组归一化的方式,避免将归一化的参数直接融合进卷积中,节省归一化操作的时间。FCOS的检测头使用了4个通道数为256的卷积作为一个分支,因此边框回归和

分类两个分支上一共有8个通道数为256的卷积,计算量非常大。为了轻量化,本实例使用了使用深度可分离卷积替换普通卷积,并且将卷积堆叠的数量从4组减少为2组。在通道数上,将256维压缩至96维,将通道数保持为8或16的倍数,这样能够享受到大部分推理框架的并行加速。最后,借鉴了yolo系列的做法,将边框回归和分类使用同一组卷积进行计算,然后分成两份。其中,预测类别的支路输出维度是数据集总类别个数K,坐标预测维度为4维的向量 (t^*, l^*, b^*, r^*) ,其中 t^*, l^*, b^*, r^* 分别为特征图P3-P7的每一个空间位置 (x, y) 的预测的真实框的坐标相对特征图的中心位置的偏移值,其预测的目标值的公式是:

$$\begin{aligned} t^* &= y - y_{\min} \\ l^* &= x - x_{\min} \\ b^* &= y_{\max} - y \\ r^* &= x_{\max} - x \end{aligned} ;$$

其中, $(y_{\min}, x_{\min}, y_{\max}, x_{\max})$ 分别表示真实标签框的四个坐标值, (t^*, l^*, b^*, r^*) 表示回归的目标值,即中心点 (x, y) 与真实框四条边界的距离。该回归方式没有利用Anchor-Based的中间媒介,由此可以达到Anchor-Free,即无锚框算法的目的。

[0033] 本实施例的其他部分与上述实施例1或2相同,故不再赘述。

[0034] 实施例4:

本实施例在上述实施例1-3任一项的基础上做进一步优化,如图2-图3所示,本实施例将主干网络ShuffleNetV2得到的C3、C4和C5进行特征融合。在YOLOV4中的PAN,都是用了步长为2的卷积进行特征图缩放,为了轻量化原则,使用 1×1 卷积来进行特征通道维度的对齐,对于上采样和下采样部分,直接使用PAN中的线性插值操作来代替步长为2的卷积,使整个特征融合模块的计算量变小。

[0035] 本实施例的其他部分与上述实施例1-3任一项相同,故不再赘述。

[0036] 实施例5:

本实施例在上述实施例1-4任一项基础上做进一步优化,本实施例中的金字塔网络中传统的特征融合模块虽然性能强大,但是堆叠的特征融合操作会使运行速度降低。为了更好的加快模型在移动端的运行速度,本发明选择了YOLO中的PAN,只有自上而下和自下而上两条通路的PAN,实现模型轻量化。

[0037] 需要说明的是,本发明使用的特征金字塔网络可以选取PAN网络,PAN网络并不唯一,其他具有相同效果特征金字塔网络的选取也适用本发明。本发明使用的主干网络可以选取ShuffleNetV2网络,ShuffleNetV2网络并不唯一,其他具有相同效果的主干网络的选取也适用本发明。无锚点检测头网络也是如此。

[0038] 本实施例的其他部分与上述实施例1-4任一项相同,故不再赘述。

[0039] 实施例6:

本实施例在上述实施例1-5任一项基础上做进一步优化,本实施例引入Anchor-free的算法作为检测器。Anchor-free的算法由于结构简单,对移动端检测更加友好,但是算法中的Centerness分支在轻量级的模型上难以训练,很难收敛,需要对其进行轻量化方向上的改进。为了确保本发明的检测模型可以用于实时目标检测,使用一阶段Anchor-free

的FCOS作为检测器,并引入GFocalLoss,去掉Centerness分支,减少了检测器的计算开销,加快检测速度。

[0040] 本实施例的其他部分与上述实施例1-5任一项相同,故不再赘述。

[0041] 实施例7:

本实施例在上述实施例1-6任一项基础上做进一步优化,本实施例使用两个通道数为96的深度可分离卷积作为每一个检测头中的定位网络,使用上述同一组卷积进行计算,然后分别进行非极大值抑制NMS操作,作为边框回归和分类的结果。其中,无锚点检测头网络的损失值GFL代表了无锚点检测头网络获取的检测特征图和多层特征图集中各图之间的误差,无锚点检测头网络的损失值GFocalLoss表示为:

$$GFL(p_{y_l}, p_{y_r}) = -|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta ((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}))$$

其中, y_l 和 y_r 表示预测模型的两个变量,这两个变量的概率值为 p_{y_l} 和 p_{y_r} , $p_{y_l} + p_{y_r} = 1$, 连续标签 y 满足 $y_l < y < y_r$, 最终预测为两个变量和概率的线性组合: $\hat{y} = y_l p_{y_l} + y_r p_{y_r}$;

绝对距离: $|y - (y_l p_{y_l} + y_r p_{y_r})|^\beta$ ($\beta \geq 0$) 是调节因子;

$((y_r - y) \log(p_{y_l}) + (y - y_l) \log(p_{y_r}))$ 是交叉熵的完全表现形式。

[0042] 本实施例的其他部分与上述实施例1-6任一项相同,故不再赘述。

[0043] 实施例8:

本实施例在上述实施例1-7任一项基础上做进一步优化,本实施例提供了一种基于无锚框算法的轻量化文字检测装置,包括采集模块、训练模块、检测模块。采集模块用于采集工业产品图像样本,并形成训练集;训练模块用于采集训练集训练网络模型,检测模块用于将待测图像输入网络模型并输出检测结果。

[0044] 训练模块就是将训练集输入进主干网络中进行特征提取得到不同尺寸的特征图,然后采用特征金字塔网络对不同尺寸的特征图进行多级预测,最后在无锚点检测头网络中进行分类和回归网络,把特征图的每个位置作为训练样本,对于特征图中的每个位置对应原图的边框都进行回归,最终得到训练好的网络模型。

[0045] 无锚点检测头网络采用了共享权重的检测头,对从特征金字塔网络出来的多尺度特征图使用同一组卷积预测检测框,然后每一层使用一个可学习的缩放值作为系数,对预测出来的框进行缩放,并且在检测头上使用了组归一化层,损失值计算方式为广义焦点损失GFL,大大减少了参数量,对提升运行速度起到了关键性作用。

[0046] 本实施例的其他部分与上述实施例1-7任一项相同,故不再赘述。

[0047] 实施例9:

本实施例在上述实施例1-8任一项基础上做进一步优化,在本实施例中,训练模块采用训练集对网络模型进行训练,预先设定训练所需的相关超参数,主干网络使用预训练权重参数,使用优化器优化总损失值,直到训练迭代次数达到最大迭代次数时停止训练。

[0048] 本实施例的其他部分与上述实施例1-8任一项相同,故不再赘述。

[0049] 实施例10:

本实施例在上述实施例1-9任一项基础上做进一步优化,如图4所示,本发明还提

供了如下方案:工业产品图像采集模块、训练和优化模块、检测模块和显示模块。

[0050] 工业产品图像采集模块主要负责包括工业产品图像以及对应标签的采集,并将采集到的数据输送到训练和优化模块;训练和优化模块包括图集确定模块、特征图确定模块、误差采集模块和模型确定模块,可接受来自工业产品图像采集模块的工业产品图像及标签,并将其用于训练和优化网络框架,主要用于训练网络模型,最终得到训练好的模型,并将其发送给检测模块;检测模块包括利用已训练好的网络模型进行检测工业产品图像,并将结果输出到显示模块;显示模块具有图像显示功能,显示检测模块发送的图像数据,得到图像是否有目标文字的检测结果。

[0051] 工业产品图像采集模块,用于获取工业产品数据集;工业产品数据集为产品上目标文字被框选的工业产品图像。

[0052] 图集确定模块,用于对工业产品数据集进行预处理得到图集。

[0053] 特征图确定模块,用于将工业产品数据集输入网络模型得到多层特征图;网络模型包括依次连接的主干网络ShuffleNetV2、特征金字塔网络和无锚点检测头;主干网络中的归一化方法包括组归一化和批标准化。

[0054] 误差计算模块,用于计算特征图与图集中各图之间的误差。

[0055] 模型确定模块,用于根据误差采用反向传播算法对网络模型进行训练得到工业产品文字检测模型,工业产品文字检测模型用于进文字检测。

[0056] 增强模块,用于对工业产品数据集进行数据增强处理得到增强过后的工业产品数据集。

[0057] 增强模块具体包括:

增强单元,用于依次对工业产品数据集进行镜像操作和尺寸调整操作,尺寸调整操作作为随机扩充或者随机裁剪。

[0058] 作为一种可选的实施方式,特征图确定模块包括:

第一特征图确定单元,用于将工业产品数据集输入残差网络得到第一特征图。

[0059] 第二特征图确定单元,用于采用特征金字塔网络对第一特征图进行处理得到第二特征图。

[0060] 第三特征图确定单元,用于将第二特征图进行检测头操作得到第三特征图。

[0061] 在实际应用中,第一特征图,具体为将工业产品数据集,可以是最初获取的工业产品数据集也可以是数据增强之后的工业产品数据集,也可以是放大或缩小后得到的操作后的工业产品数据集,输入进主干网络中最后3层(第3,4,5层)输出的特征图为第一特征图,其中主干网络中的批标准化操作可以替换为批组标准化。

[0062] 在实际应用中,特征金字塔网络从主干网络的第3阶段开始,总共分为三个步骤,具体方案如下:

首先:获取ShuffleNetV2网络生成的第3,4,5层的特征图。

[0063] 其次:将第3层的特征图经过一个 3×3 的卷积核进一步单独的提取特征,并将其维度降为与第4层的特征图一致,将第5层的特征图通过上采样使其与第4层的特征图维度一致。

[0064] 最后:将第4层的特征图进行拷贝,将其与步骤52生成的两个特征层进行融合,得到新的第5层特征图。

[0065] 以与特征金字塔网络从主干网络的第3阶段开始的三个步骤相类似的方式获得新的第3层特征图(将第3层、第4层和第5层的特征图均进行上采样得到与第2层特征图的维度相同的图,将这三个特征图进行融合得到新的第3层特征图)和新的第4层特征图(将第4层和第5层的特征图进行上采样得到与第3层特征图的维度相同的图,将两个图与第3层的特征图进行融合得到新的第4层特征层),并将新的第3层特征层、第4层特征层和第5层特征层再经过类似于步骤52和步骤53的步骤得到维度大小为 $[256, h/n, w/n]$ 的第二特征图,第二特征图的维度大小与工业产品数据集预处理后三个图的大小一致。

[0066] 本发明提供的基于无锚框算法的轻量化文字检测装置被具体实现为一种计算机程序产品,计算机程序产品可以包括计算机可读存储介质,其上载有用于执行本发明提供的基于无锚框算法的轻量化文字检测装置的计算机可读程序指令。

[0067] 计算机可读存储介质可以是保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是但不限于电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意组合。

[0068] 本实施例的其他部分与上述实施例1-9任一项相同,故不再赘述。

[0069] 最后应说明的是,以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换,而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

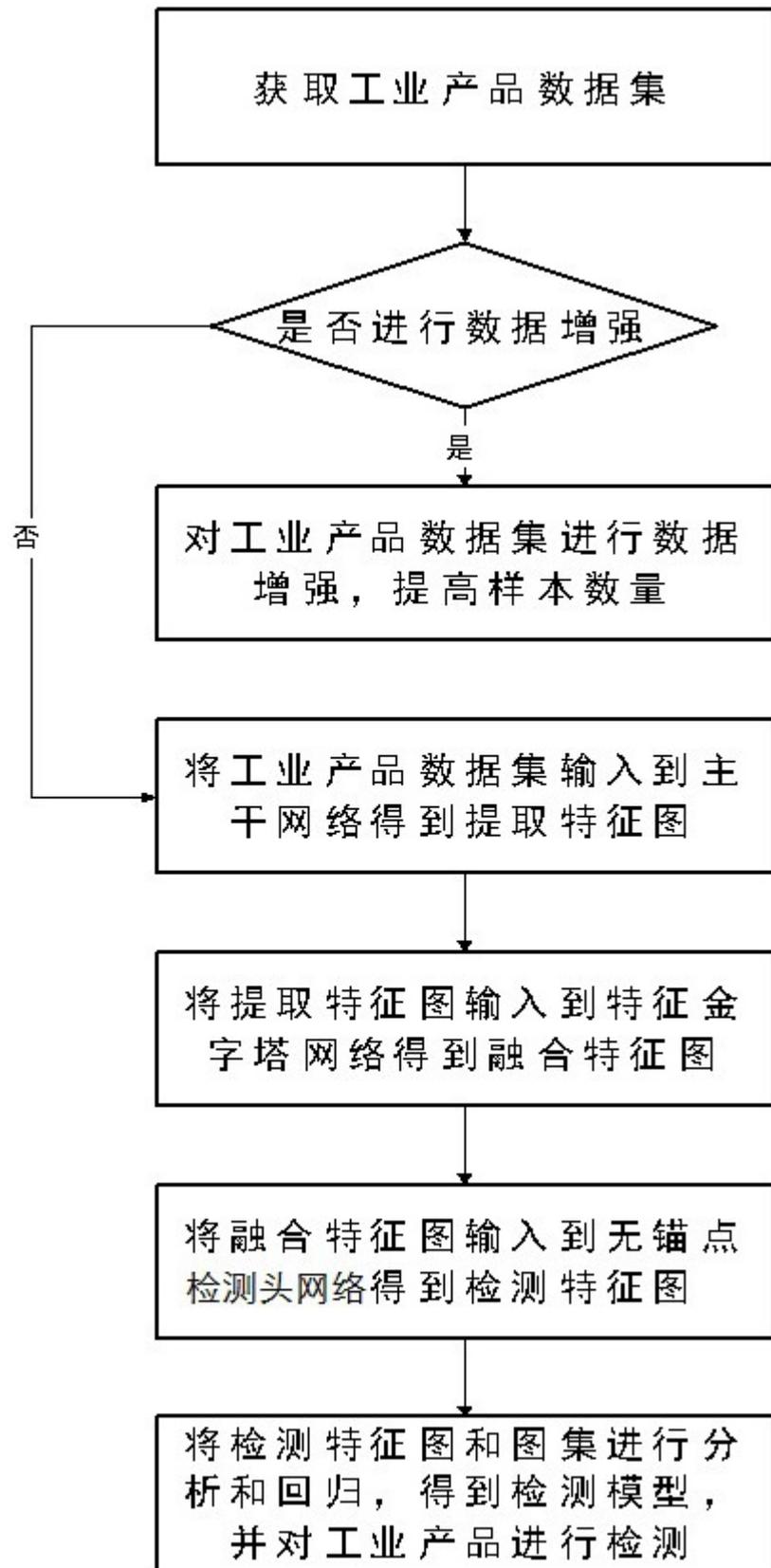


图1

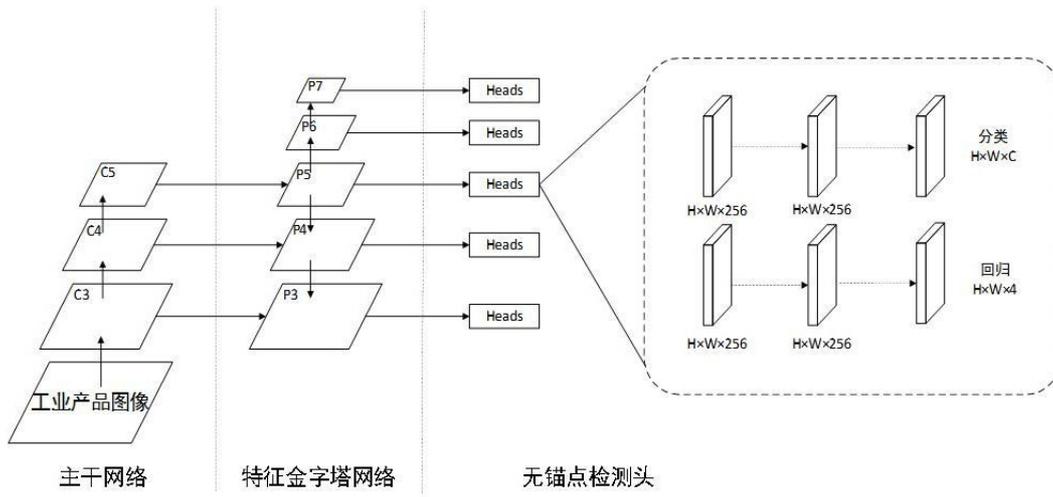


图2

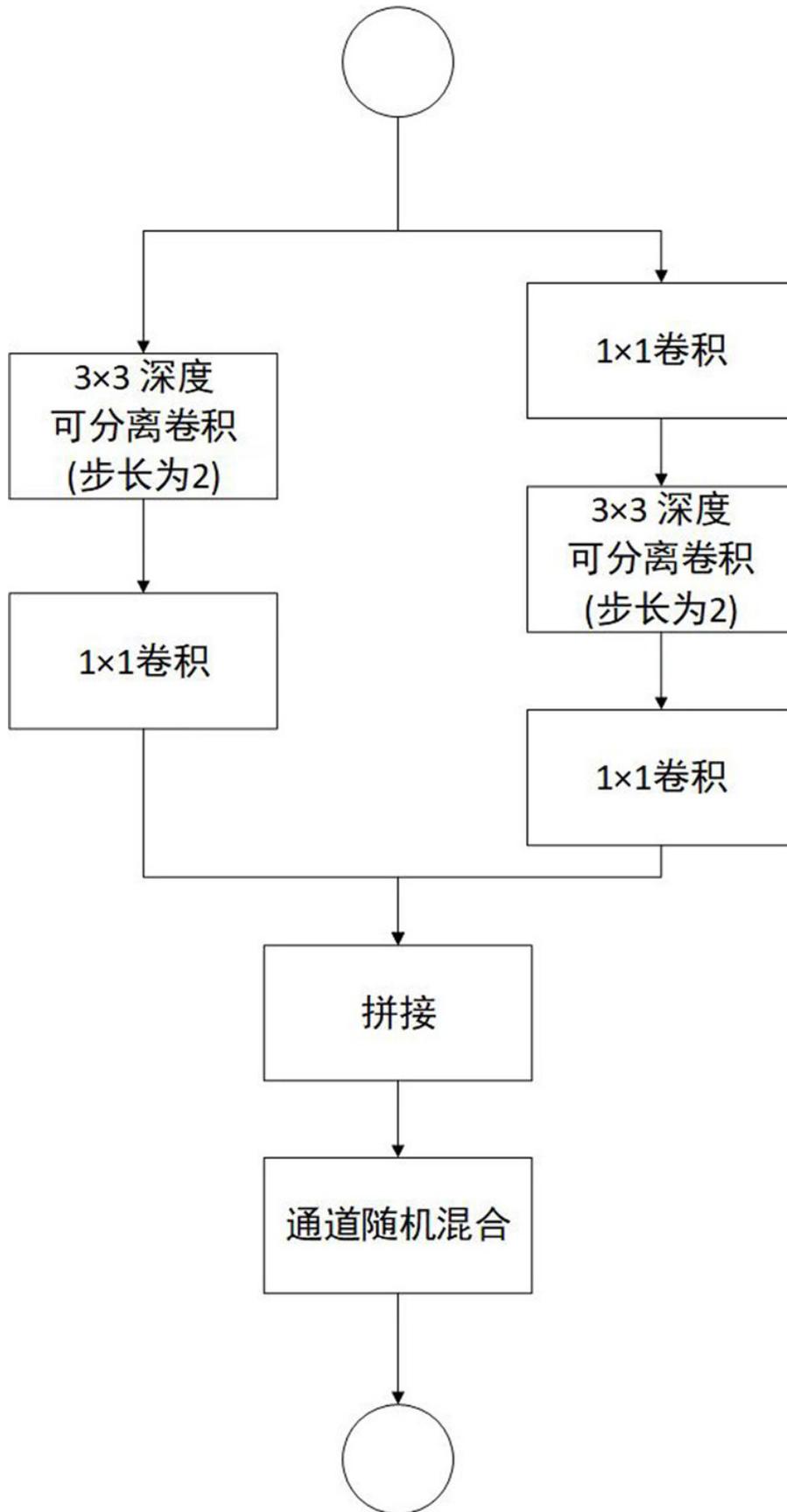


图3

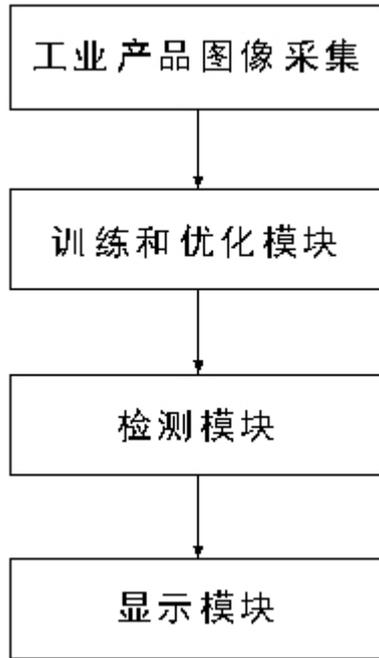


图4