



(12) 发明专利

(10) 授权公告号 CN 112364726 B

(45) 授权公告日 2024. 06. 04

(21) 申请号 202011163480.5

(22) 申请日 2020.10.27

(65) 同一申请的已公布的文献号

申请公布号 CN 112364726 A

(43) 申请公布日 2021.02.12

(73) 专利权人 重庆大学

地址 400044 重庆市沙坪坝区沙正街174号

(72) 发明人 唐倩 冯琪翔 李代杨 张志豪

(74) 专利代理机构 重庆谢成律师事务所 50224

专利代理师 谢殿武

(51) Int. Cl.

G06V 30/14 (2022.01)

G06V 30/40 (2022.01)

G06V 10/82 (2022.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 107133616 A, 2017.09.05

CN 110287960 A, 2019.09.27

CN 111310760 A, 2020.06.19

CN 111583180 A, 2020.08.25

US 2016191194 A1, 2016.06.30

US 2018359029 A1, 2018.12.13

US 7143153 B1, 2006.11.28

WO 2020164281 A1, 2020.08.20

WO 9717988 A1, 1997.05.22

李代杨. 基于深度神经网络的汽车刹车片喷码检测识别技术研究. 万方硕士学位论文. 2021, 1-7.

LEW, S. Environmental controls on the abundance of methanotrophs and methanogens in peat bog lakes. SCIENCE OF THE TOTAL ENVIRONMENT. 2018, 1201-1211.

唐倩文; 陈良育. 基于复杂网络理论的Java开源系统演化分析. 计算机科学. 2018, (第08期), 173-180.

审查员 张驰

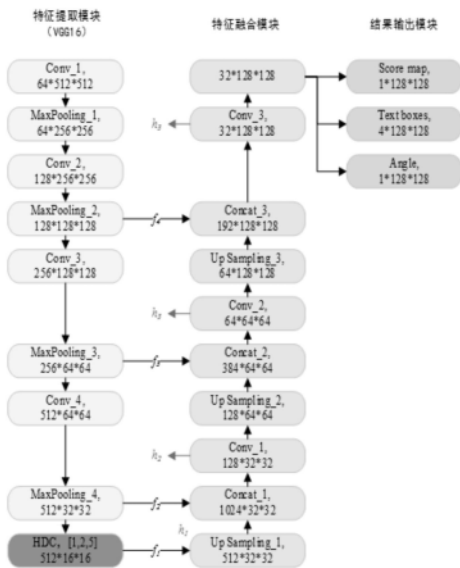
权利要求书2页 说明书5页 附图2页

(54) 发明名称

基于改进EAST的零件喷码字符定位的方法

(57) 摘要

本发明提供一种基于改进EAST的零件喷码字符定位的方法, 所述方法以现有EAST网络结构为基础, 改进现有EAST网络的网络结构、标签生成、损失函数和候选区域的处理, 获得改进后的网络结构、改进后的标签生成、改进后的损失函数和候选区域后处理, 并用改进后的EAST网络来定位字符。本申请在现有EAST算法的基础上就其网络结构、标签生成、损失函数和候选区域后处理进行优化, 改进后的深度学习网络可有效扩大感受视野, 提升对长文本的识别效率, 实现对零件喷码字符的准确定位。



1. 一种基于改进EAST的零件喷码字符定位的方法,其特征在于:所述方法以现有EAST网络结构为基础,改进现有EAST网络的网络结构、标签生成、损失函数和候选区域的处理,获得改进后的网络结构、改进后的标签生成、改进后的损失函数和候选区域后处理,并用改进后的EAST网络来定位字符;

所述改进现有EAST网络的网络结构具体包括:

构建VGG16模型,所述VGG16模型包括5个卷积层和5个池化层,卷积核大小为 3×3 ,步长均为1;

改进VGG16模型,将空洞系数为[1,2,5]的混合空洞卷积HDC替换VGG16网络结构中最后一个阶段的卷积操作;

所述改进标签生成具体包括:采用长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形;

所述改进损失函数具体包括:采用dice loss损失函数替代EAST算法中的平衡交叉熵损失函数;

所述改进候选区域的处理具体包括:采用了基于预排序的局部感知NMS算法,根据候选框的左上角坐标的值对所有候选框进行排序;

所述改进标签生成具体包括如下步骤:

以长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形;

所述四边形的短边在原标签备注的短边长度的基础上缩小0.3倍;

所述四边形的长边在原标签备注的长边长度的基础上缩小shrunk_rate倍数:

$$shrunk_rate = 0.3 \times \frac{length_s}{length_l} \quad (1)$$

其中,shrunk_rate表示长边的缩小倍数,length_s表示四边形短边的长度,length_l表示四边形长边的长度;

所述dice loss损失函数采用如下方法确定:

$$L = L_s + \lambda_g L_g = L_s + \lambda_g (L_{IoU} + \lambda_\theta L_\theta) \quad (2)$$

其中,L表示改进后的dice loss损失函数,L_s表示得分图损失,L_g表示几何损失,λ_g表示几何损失的系数,L_{IoU}表示重叠面积损失,λ_θ表示角度损失的系数,L_θ表示角度损失;

所述改进候选区域的处理具体包括基于预排序的局部感知NMS算法,所述基于预排序的局部感知NMS算法具体步骤如下:

S1:初始化所有候选矩形框的集合boxes、重叠度阈值λ和激变阈值ω;

S2:提取boxes中的所有矩形框的左上角坐标(x₁,y₁),分别按照x₁和y₁从小到大的顺序排列,根据ω分别计算并比较x₁和y₁的激变次数,将boxes按照激变次数从小到大的顺序排列;

S3:设为S空集,p为空集,对于g∈boxes,其中g表示所有候选矩形框的集合boxes中的元素,即候选矩形框,按照boxes中的所有矩形框的激变次数从小到大的顺序依次执行:如果p不为空且p和g重叠面积比例大于λ,则p为p和g合并后的矩形框;如果p不为空且p和g重叠面积比例不大于λ,则将p放入S;如果p为空,则p=g;

S4:如果p不为空,则将p放入S;

S5:输出集合S,所述S表示筛选后的矩形框集合;

所述改进候选区域的处理还包括候选框合并,所述候选框合并采用如下方法确定:

$$S(a) = \eta_g S(g) + \eta_p S(p) \quad (3)$$

$$a_i = \eta_g S(g) g_i + \eta_p S(p) p_i \quad (4)$$

其中,a为合并后的矩形框,g和p表示满足重叠条件的候选矩形框,i表示矩形框中的参数索引, $i=1,2,3\cdots 9$,4个顶点共8个坐标,1个角度值,S表示该矩形框的得分值, η_g 表示候选矩形框g的面积系数, η_p 表示候选矩形框p的面积系数;

所述面积系数 η_g 和 η_p 采用如下方法设置:

判断候选矩形框g的面积是否大于候选矩形框p的面积,若是,则 $\eta_g=1.1$, $\eta_p=1$,若否,则 $\eta_g=1$, $\eta_p=1.1$ 。

基于改进EAST的零件喷码字符定位的方法

技术领域

[0001] 本发明涉及字符识别技术领域,尤其涉及一种基于改进EAST的零件喷码字符定位的方法。

背景技术

[0002] 随着信息科学技术的发展,图像处理和机器视觉技术已经在工业领域得到了广泛应用。字符文本检测作为使用最为广泛的机器视觉技术,其工业应用一直是研究的重点与难点。而字符文本检测的首要步骤就是对字符文本区域进行定位。EAST算法作为优秀的自然场景文本定位算法之一,在多个公开数据集中有着出色的表现。但是受其感受野大小、训练过程设置等方面的限制,EAST算法在字符定位中表现不够理想,且其对长文本的检测效果有待提高。

[0003] 因此,亟需一种对字符分割的依赖性要小、通信性高的端到端字符识别方法。

发明内容

[0004] 有鉴于此,本发明提供一种基于改进EAST的零件喷码字符定位的方法,其特征在于:所述方法以现有EAST网络结构为基础,改进现有EAST网络的网络结构、标签生成、损失函数和候选区域的处理,获得改进后的网络结构、改进后的标签生成、改进后的损失函数和候选区域后处理,并用改进后的EAST网络来定位字符;

[0005] 所述改进现有EAST网络的网络结构具体包括:

[0006] 构建VGG16模型,所述VGG16模型包括5个卷积层和5个池化层,卷积核大小为 3×3 ,步长均为1;

[0007] 改进VGG16模型,将空洞系数为[1,2,5]的混合空洞卷积HDC替换VGG16网络结构中最后一个阶段的卷积操作;

[0008] 所述改进标签生成具体包括:采用长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形;

[0009] 所述改进损失函数具体包括:采用dice loss损失函数替代EAST算法中的平衡交叉熵损失函数;

[0010] 所述改进候选区域的处理具体包括:采用了基于预排序的局部感知NMS算法,根据候选框的左上角坐标的值对所有候选框进行排序。

[0011] 进一步,所述改进标签生成具体包括如下步骤:

[0012] 以长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形;

[0013] 所述四边形的短边在原标签备注的短边长度的基础上缩小0.3倍;

[0014] 所述四边形的长边在原标签备注的长边长度的基础上缩小shrunk_rate倍数:

$$[0015] \quad shrunk_rate = 0.3 \times \frac{length_s}{length_l} \quad (1)$$

[0016] 其中,shrunk_rate表示长边的缩小倍数,length_s表示四边形短边的长度,length_l表示四边形长边的长度。

[0017] 进一步,所述dice loss损失函数采用如下方法确定:

$$[0018] \quad L = L_s + \lambda_g L_g = L_s + \lambda_g (L_{IoU} + \lambda_\theta L_\theta) \quad (2)$$

[0019] 其中,L表示改进后的dice loss损失函数,L_s表示得分图损失,L_g表示几何损失,λ_g表示几何损失的系数,L_{IoU}表示重叠面积损失,λ_θ表示角度损失的系数,L_θ表示角度损失。

[0020] 进一步,所述改进候选区域的处理具体包括基于预排序的局部感知NMS算法,所述基于预排序的局部感知NMS算法具体步骤如下:

[0021] S1:初始化所有候选矩形框的集合boxes、重叠度阈值λ和激变阈值ω;

[0022] S2:提取boxes中的所有矩形框的左上角坐标(x₁,y₁),分别按照x₁和y₁从小到大的顺序排列,根据ω分别计算并比较x₁和y₁的激变次数,将boxes按照激变次数从小到大的顺序排列;

[0023] S3:设为s空集,p为空集,对于g∈boxes,其中g表示所有候选矩形框的集合boxes中的元素,即候选矩形框,按照boxes中的所有矩形框的激变次数从小到大的顺序依次执行:如果p不为空且p和g重叠面积比例大于λ,则p为p和g合并后的矩形框;如果p不为空且p和g重叠面积比例不大于λ,则将p放入s;如果p为空,则p=g;

[0024] S4:如果p不为空,则将p放入s;

[0025] S5:输出集合S,所述S表示筛选后的矩形框集合。

[0026] 进一步,所述改进候选区域的处理还包括候选框合并,所述候选框合并采用如下方法确定:

$$[0027] \quad S(a) = \eta_g S(g) + \eta_p S(p) \quad (3)$$

$$[0028] \quad a_i = \eta_g S(g) g_i + \eta_p S(p) p_i \quad (4)$$

[0029] 其中,a为合并后的矩形框,g和p表示满足重叠条件的候选矩形框,i表示矩形框中的参数索引,i=1,2,3...9,4个顶点共8个坐标,1个角度值,S表示该矩形框的得分值,η_g表示候选矩形框g的面积系数,η_p表示候选矩形框p的面积系数;

[0030] 所述面积系数η_g和η_p采用如下方法设置:

[0031] 判断候选矩形框g的面积是否大于候选矩形框p的面积,若是,则η_g=1.1,η_p=1,若否,则η_g=1,η_p=1.1。

[0032] 本发明的有益技术效果:本申请在现有EAST算法的基础上就其网络结构、标签生成、损失函数和候选区域后处理进行优化,改进后的深度学习网络可有效扩大感受视野,提升对长文本的识别效率,实现对零件喷码字符的准确定位。

附图说明

[0033] 下面结合附图和实施例对本发明作进一步描述:

[0034] 图1是本发明提出的改进的EAST算法网络结构的示意图。

[0035] 图2是本发明标签生成过程示意图。

[0036] 图3是本发明检测效果与原算法检测效果对比图。

具体实施方式

[0037] 以下结合说明书附图对本发明做出进一步的说明：

[0038] 本发明提供一种基于改进EAST的零件喷码字符定位的方法，其特征在于：所述方法以现有EAST网络结构为基础，改进现有EAST网络的网络结构、标签生成、损失函数和候选区域的处理，获得改进后的网络结构、改进后的标签生成、改进后的损失函数和候选区域后处理，并用改进后的EAST网络来定位字符；

[0039] 所述改进现有EAST网络的网络结构具体包括：

[0040] 构建VGG16模型，所述VGG16模型包括5个卷积层和5个池化层，卷积核大小为 3×3 ，步长均为1；

[0041] 改进VGG16模型，将空洞系数为[1, 2, 5]的混合空洞卷积HDC替换VGG16网络结构中最后一个阶段的卷积操作；

[0042] 所述改进标签生成具体包括：采用长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形；

[0043] 所述改进损失函数具体包括：采用dice loss损失函数替代EAST算法中的平衡交叉熵损失函数；

[0044] 所述改进候选区域的处理具体包括：采用了基于预排序的局部感知NMS算法，根据候选框的左上角坐标的值对所有候选框进行排序。

[0045] 在卷积神经网络中，感受野的大小是由卷积核、卷积步长等参数共同决定的。在标准的VGG16模型中，是通过卷积层的叠加和池化操作来实现感受野的扩大。但是这种扩大感受野的方法由于受卷积层数的限制，效果很有限，扩大卷积核的大小又会导致参数过多，而通过池化层增加感受野则会导致信息的丢失。这些因素使得VGG16网络感受野的大小受到限制。空洞卷积是一种卷积核大小本身不变，但是加入许多空洞来将整个卷积矩阵变大，扩展卷积区域，从而达到既不增加参数又扩大感受野的方法，但会出现网格效应和小目标无法得到检测的现象。但有人提出了混合空洞卷积方案，认为只要在搭建神经网络过程中卷积系数设置满足以下三个要求就可以避免上述的问题，按照这样的思路，对VGG16最后一个阶段的卷积操作(Conv_5)进行替换，将原来的连续3次的标准卷积替换为空洞系数为[1, 2, 5]的混合空洞卷积(HDC)，使原来该部分 7×7 的感受野扩大为 17×17 ，从而得到感受野更大的VGG16网络。新的网络结构如图1所示，深蓝色部分即为本文对VGG16网络改进部分，其中Maxpooling表示最大池化操作，Up Sampling表示上采样操作，Concat表示特征图拼接。

[0046] 上述技术方案在现有EAST算法的基础上就其网络结构、标签生成、损失函数和候选区域后处理进行优化，改进后的深度学习网络可有效扩大感受视野，提升对长文本的识别效率，实现对零件喷码字符的准确的定位。

[0047] 在本实施例中，所述改进标签生成具体包括如下步骤：

[0048] 以长短边区别对待的方式缩小目标训练图像的字符区域形成的四边形；

[0049] 所述四边形的短边在原标签备注的短边长度的基础上缩小0.3倍；

[0050] 所述四边形的长边在原标签备注的长边长度的基础上缩小shrunk_rate倍数：

$$[0051] \quad shrunk_rate = 0.3 \times \frac{length_s}{length_l} \quad (1)$$

[0052] 其中,shrunk_rate表示长边的缩小倍数,length_s表示四边形短边的长度,length_l表示四边形长边的长度。

[0053] 在网络训练过程中,所有训练图像的字符区域标签均为四边形,其四个点的坐标以(x₁,y₁,x₂,y₂,x₃,y₃,x₄,y₄)的格式表示。x和y分别表示点的横坐标和纵坐标。下标“1”、“2”、“3”、“4”分别表示左上、右上、右下和左下角的点,即顺时针顺序标注。在EAST算法训练时,采用的是RBOX的形式来生成训练标签,即5个通道{d₁,d₂,d₃,d₄,θ},其中d_i表示训练标签内的点到标签四条边的距离(图1中的text boxes),θ表示旋转角度(图1中的angle)。另有1个通道来记录图像中各点是字符点的可能性评分(图1中的score map)。为了减少标签标注误差带来的干涉,EAST算法中将得分图(score map)在原标签标注的最小外接矩形的基础上缩小了0.3倍。整个标签生成过程如图2所示。这种将标签矩形缩小的方法可以减少字符边缘区域带来的干扰,提高字符定位的准确度。但是从图2中可以明显看出,由于字符区域的长边基数较大,缩小0.3倍会导致很大面积的字符区域没有进入得分图中。这样的训练对长字符定位来说是很糟糕的。针对这个问题,对EAST算法的标签训练进行改进,采用长短边区别对待的方式,对于标签矩形的短边仍然按照0.3倍进行缩小,而长边的缩小倍数为0.3倍的矩形短边长度与矩形长边长度之比。

[0054] 在本实施例中,所述dice loss损失函数采用如下方法确定:

$$[0055] \quad L = L_s + \lambda_g L_g = L_s + \lambda_g (L_{IoU} + \lambda_\theta L_\theta) \quad (2)$$

[0056] 其中,L表示改进后的dice loss损失函数,L_s表示得分图损失,L_g表示几何损失,λ_g表示几何损失的系数,L_{IoU}表示重叠面积损失,λ_θ表示角度损失的系数,L_θ表示角度损失。

[0057] 对于神经网络训练而言,损失函数就是整个网络训练的优化目标,损失函数设置的好坏直接影响网络的训练结果和最终参数的检测效果。对于本发明提出的FCN网络而言,损失函数L主要由得分图损失L_s和几何损失L_g两部分组成,其中几何损失L_g又是由重叠面积损失L_{IoU}和角度损失L_θ组成。对于得分图损失L_s,其本质上是一个二分类的损失函数,即字符区域为正样本,其他区域为负样本,且其正负样本非常不平衡,字符区域面积远小于其他区域。EAST算法采用了平衡交叉熵损失函数来解决样本不平衡问题,而近几年的研究发现dice loss在应对样本不平衡问题时有着更为优秀的表现。因此,本发明采用dice loss来表达得分图损失,较之EAST算法可以很好应对样本不平衡问题。

[0058] 在本实施例中,所述改进候选区域的处理具体包括基于预排序的局部感知NMS算法,所述基于预排序的局部感知NMS算法具体步骤如下:

[0059] S1:初始化所有候选矩形框的集合boxes、重叠度阈值λ和激变阈值ω;

[0060] S2:提取boxes中的所有矩形框的左上角坐标(x₁,y₁),分别按照x₁和y₁从小到大的顺序排列,根据ω分别计算并比较x₁和y₁的激变次数,将boxes按照激变次数从小到大的顺序排列;

[0061] S3:设为S空集,p为空集,对于g∈boxes,其中g表示所有候选矩形框的集合boxes中的元素,即候选矩形框,按照boxes中的所有矩形框的激变次数从小到大的顺序依次执行:如果p不为空且p和g重叠面积比例大于λ,则p为p和g合并后的矩形框;如果p不为空且p和g重叠面积比例不大于λ,则将p放入S;如果p为空,则p=g;

[0062] S5:输出集合S,所述S表示筛选后的矩形框集合。

[0063] 通常情况下,使用该网络进行字符区域检测时,通过FCN提取出来的候选框可能有

数百甚至上千个。它们大多是相互重叠的,因此将这数量巨大的候选框进行合并就是整个字符区域定位算法的最后一个步骤了。以往的后续框合并往往采用非极大值抑制,即NMS算法,但是这种算法的时间复杂度为 $O(n^2)$,非常耗时。针对这种情况,EAST算法采用了局部感知的NMS算法。该算法只比较排序相邻的两个候选区域,重叠度大于阈值则合并,小于则保留。其时间复杂度取决于候选区域的排列情况,最好的情况为 $O(n)$ (即应该合并的候选框都排列在一起),最坏的情况为 $O(n^2)$ (互不重叠的候选框交替排列)。同时,对于满足重叠度条件的两个候选框不再按照标准NMS那样舍弃一个留下另一个,而是对两个候选框根据其得分情况进行合并。

[0064] 可以看出,局部感知的NMS算法的时间复杂度依赖于随机的候选区排列情况。针对整个问题,本发明对局部感知NMS算法进行了改进,提出基于预排序的局部感知NMS算法,根据候选框的左上角坐标的值对所有候选框进行排序,这样就可以将所有可能进行合并的候选框都排列在一起,达到时间复杂度尽可能为 $O(n)$ 的效果。同时,将可能合并的候选框排列在一起也有利于提高合并的精度,从而改善最后的文本定位效果。

[0065] 在本实施例中,所述改进候选区域的处理还包括候选框合并,所述候选框合并采用如下方法确定:

$$[0066] \quad S(a) = \eta_g S(g) + \eta_p S(p) \quad (3)$$

$$[0067] \quad a_i = \eta_g S(g) g_i + \eta_p S(p) p_i \quad (4)$$

[0068] 其中, a 为合并后的矩形框, g 和 p 表示满足重叠条件的候选矩形框, i 表示矩形框中的参数索引, $i=1,2,3\cdots 9$,4个顶点共8个坐标,1个角度值, S 表示该矩形框的得分值, η_g 表示候选矩形框 g 的面积系数, η_p 表示候选矩形框 p 的面积系数;

[0069] 所述面积系数 η_g 和 η_p 采用如下方法设置:

[0070] 判断候选矩形框 g 的面积是否大于候选矩形框 p 的面积,若是,则 $\eta_g=1.1$, $\eta_p=1$,若否,则 $\eta_g=1$, $\eta_p=1.1$ 。

[0071] 此外,在原来的局部感知NMS算法中,对候选矩形框进行合并时,也只考虑得分情况,对于长文本而言,候选框的面积信息也同样重要,面积更大的候选框涵盖整个文本区域的可能性就更大。因此,本文在候选框合并中,引入了面积系数 η , g 和 p 中面积较大者的 η 设为1.05,另一个为1。

[0072] 按照以上基于EAST的改进字符定位算法实施,对比算法改进前后的检测效果,如图3所示。可以明显看出原EAST算法对于长文本检测的缺陷,第一行字符最右侧的矩形框没有覆盖到所有字符区域,第二行字符的最后一个“1”也只被覆盖了一半。而针对长文本检测使用了优化的本发明算法,则很好地覆盖了整个字符区域,没有出现漏掉字符区域首尾的情况。

[0073] 最后说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或者等同替换,而不脱离本发明技术方案的宗旨和范围,其均应涵盖在本发明的权利要求范围当中。

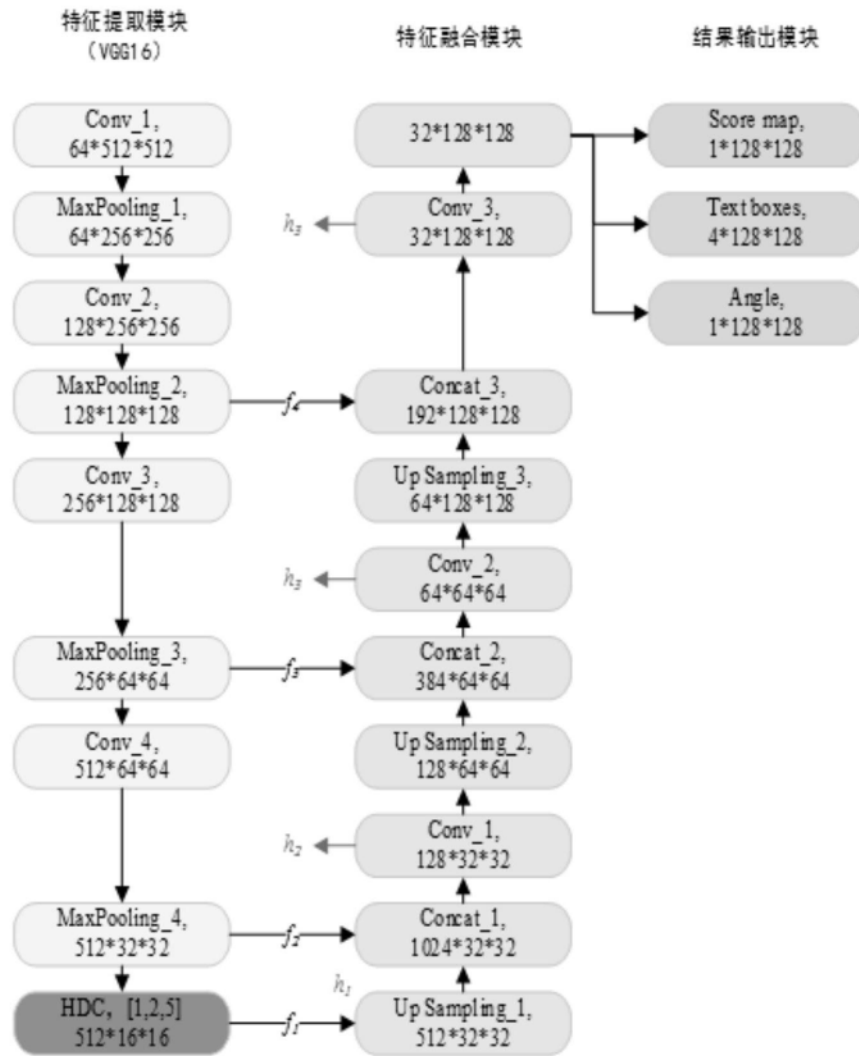


图1

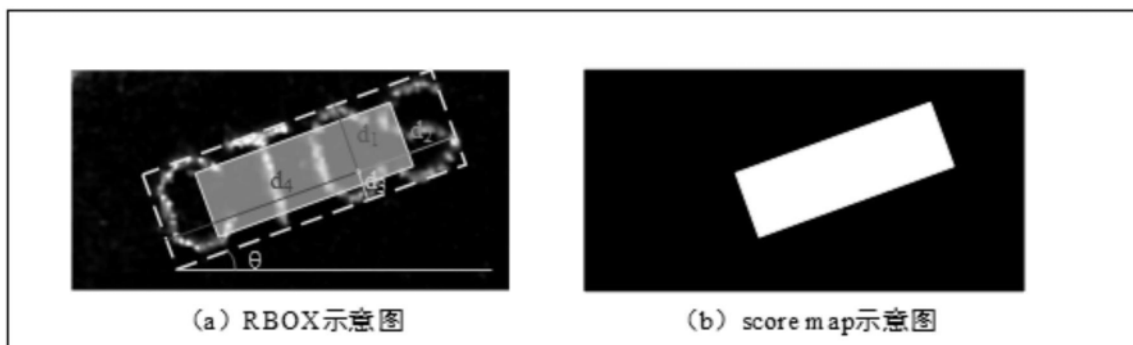


图2

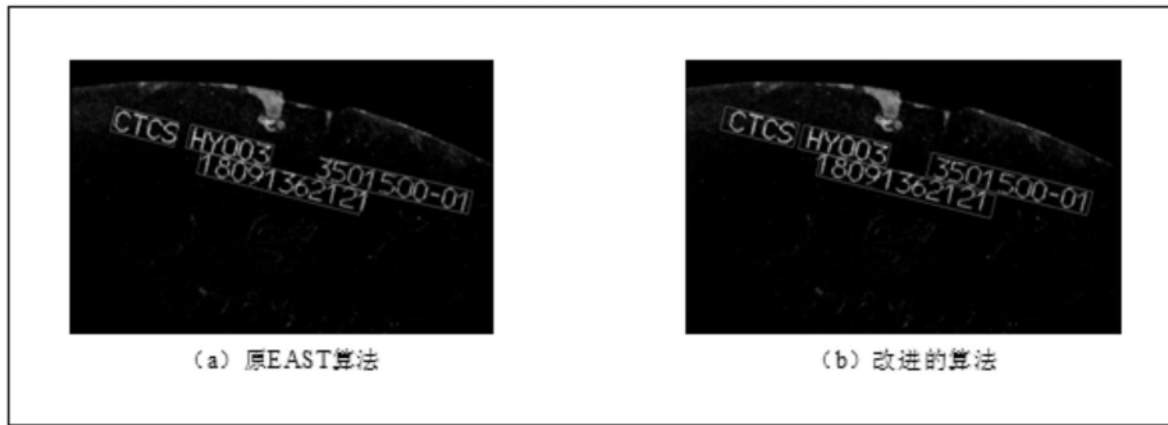


图3