



(12) 发明专利申请

(10) 申请公布号 CN 116452958 A

(43) 申请公布日 2023. 07. 18

(21) 申请号 202310283937.3

G06N 3/0464 (2023.01)

(22) 申请日 2023.03.22

G06N 3/08 (2023.01)

(71) 申请人 吉林大学

G06V 10/26 (2022.01)

地址 130000 吉林省长春市前进大街2699号

G06V 20/70 (2022.01)

(72) 发明人 贾继伟 孟祥全 詹中涵

(74) 专利代理机构 深圳国联专利代理事务所
(特殊普通合伙) 44465

专利代理师 宿晓燕

(51) Int. Cl.

G06V 20/00 (2022.01)

G06V 10/82 (2022.01)

G06V 10/774 (2022.01)

G06V 10/764 (2022.01)

G06V 10/40 (2022.01)

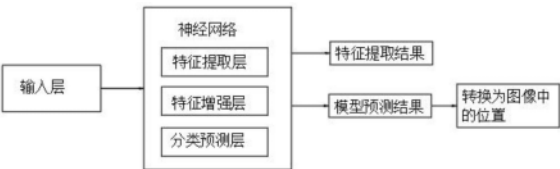
权利要求书1页 说明书4页 附图3页

(54) 发明名称

一种基于深度学习的直线目标体识别方法

(57) 摘要

本发明提供了一种基于深度学习的直线目标体识别方法,包括以下步骤:S1:对图像、图片或视频构建数据集;S2:以车道线检测算法为基本框架,参照车道线检测算法中的网络结构和损失函数,搭建神经网络;S3:使用数据集对神经网络进行训练;S4:对训练结果进行后处理,将结果转化进行位置识别;S5:对训练好的网络进行评估;计算预测成功的比例作为准确率。本发明中,鉴于2D图像中的直线目标体物理特点,以及其在图像中的位置特点,选择深度学习的方法,使用卷积神经网络进行识别;参考车道线检测算法,以ResNet作为特征提取网络,训练时加入用于特征增强的网络层,提高特征提取效果,进而提高识别效果。



1. 一种基于深度学习的直线目标体识别方法,其特征在于,包括以下步骤:

S1:对图像、图片或视频构建数据集;

S2:以车道线检测算法为基本框架,参照车道线检测算法中的网络结构和损失函数,搭建神经网络;

S3:使用数据集对神经网络进行训练;

S4:对训练结果进行后处理,将结果转化进行位置识别;

S5:对训练好的网络进行评估;计算预测成功的比例作为准确率。

2. 根据权利要求1所述的一种基于深度学习的直线目标体识别方法,其特征在于,在S5中,阈值为10像素。

3. 根据权利要求1所述的一种基于深度学习的直线目标体识别方法,其特征在于,所述评估结果的判定方法为:在每一个水平行上,预测点与真实标记点的距离值小于阈值,则判定预测成功;反之,判定预测失败。

4. 根据权利要求1所述的一种基于深度学习的直线目标体识别方法,其特征在于,包括输入层;神经网络包括特征提取层、特征增强层和分类预测层;

输入层,用于获取图片、图像及视频;

特征提取层,用于图片特征提取,并在18层残差网络的基础上,对网络后三层的输出进行提取,构建网络层进行特征增强;

特征增强层,为包括卷积、归一化及Relu的conv_bn_relu块,以对残差网络的三个输出,分别构建由多个conv_bn_relu块组成的网络,最后对三个结果进行合并,对合并后的结果再次进行多个conv_bn_relu块操作以及卷积操作,形成特征增强网络层,用于图片特征增强处理;

分类预测层,包括全连接层,用于生成预测图。

5. 根据权利要求4所述的一种基于深度学习的直线目标体识别方法,其特征在于,所述特征提取层包括整体网络的损失函数,损失函数包括整体分类损失、整体预测损失、分割损失以及特征增强层损失,上述损失均采用交叉熵损失函数。

6. 根据权利要求4所述的一种基于深度学习的直线目标体识别方法,其特征在于,对于图像中的直线目标体,利用神经网络将识别任务视为语义分割任务。

一种基于深度学习的直线目标体识别方法

技术领域

[0001] 本发明属于计算机视觉、目标检测技术领域,涉及一种基于深度学习的直线目标体识别方法。

背景技术

[0002] 目前,利用神经网络进行目标检测主流的方法是采用卷积神经网络,基于卷积网络搭建的许多网络结构已经在目标检测等诸多领域取得了很好的效果及广泛的应用。

[0003] 残差网络(ResNet)是一种卷积神经网络,其最初提出是为了解决网络深度加深反而使模型结果变差的情况。残差网络通过利用残差结构,可以很好的解决这一问题。残差网络有效的增加了网络深度,提升了准确率,是目前为止应用最为广泛的卷积神经网络之一。

[0004] 针对于2D目标检测,在现有的检测算法中,基于图像的方法无法解决复杂的场景问题,特别是模糊,遮挡等问题;同时,用卷积网络提取特征后直接进行预测,特征提取的效果不是很好,使最终预测效果也不是很好。

[0005] 本发明参考类比Qin,ZequnandWang,HuanyuandLi,Xi提出的车道线检测算法;我们参考了相关研究更多的车道线检测算法;车道线可以看作是直线目标体的一种,并认为车道线检测算法可以推广到一般性问题上。

发明内容

[0006] 本发明的目的是针对现有的技术存在上述问题,提出了一种基于深度学习的直线目标体识别方法。

[0007] 本发明的目的可通过下列技术方案来实现:

[0008] 一种基于深度学习的直线目标体识别方法,包括以下步骤:

[0009] S1:对图像、图片或视频构建数据集;

[0010] S2:以车道线检测算法为基本框架,参照车道线检测算法中的网络结构和损失函数,搭建神经网络;

[0011] S3:使用数据集对神经网络进行训练;

[0012] S4:对训练结果进行后处理,将结果转化进行位置识别;

[0013] S5:对训练好的网络进行评估;计算预测成功的比例作为准确率。

[0014] 进一步的,在S5中,阈值为10像素。

[0015] 进一步的,所述评估结果的判定方法为:在每一个水平行上,预测点与真实标记点的距离值小于阈值,则判定预测成功;反之,判定预测失败。

[0016] 进一步的,包括输入层;神经网络包括特征提取层、特征增强层和分类预测层;

[0017] 输入层,用于获取图片、图像及视频;

[0018] 特征提取层,用于图片特征提取,并在18层残差网络的基础上,对网络后三层的输出进行提取,构建网络层进行特征增强;

[0019] 特征增强层,为包括卷积、归一化及Relu的conv_bn_relu块,以对残差网络的三个

输出,分别构建由多个conv_bn_relu块组成的网络,最后对三个结果进行合并,对合并后的结果再次进行多个conv_bn_relu块操作以及卷积操作,形成特征增强网络层,用于图片特征增强处理;

[0020] 分类预测层,包括全连接层,用于生成预测图。

[0021] 进一步的,所述特征提取层包括整体网络的损失函数,损失函数包括整体分类损失、整体预测损失、分割损失以及特征增强层损失,上述损失均采用交叉熵损失函数。

[0022] 进一步的,对于图像中的直线目标体,利用神经网络将识别任务视为语义分割任务。

[0023] 与现有技术相比,本发明具有以下优点:鉴于2D图像中的直线目标体物理特点,以及其在图像中的位置特点,选择深度学习的方法,使用卷积神经网络进行识别;参考车道线检测算法,以ResNet作为特征提取网络,训练时加入用于特征增强的网络层,提高特征提取效果,进而提高识别效果。

附图说明

[0024] 图1是本发明的系统示意图。

[0025] 图2是本发明中实施例2训练分类误差示意图。

[0026] 图3是本发明中实施例2训练特征增强层训练误差示意图。

[0027] 图4是本发明中实施例2验证集均交并比示意图。

[0028] 图5是本发明中实施例2top3准确率示意图。

具体实施方式

[0029] 以下是本发明的具体实施例并结合附图,对本发明的技术方案作进一步的描述,但本发明并不限于这些实施例。

[0030] 如图1所示,本实施例提供一种基于深度学习的直线目标体识别方法,包括以下步骤:

[0031] S1:对图像、图片或视频构建数据集;

[0032] S2:以车道线检测算法为基本框架,参照车道线检测算法中的网络结构和损失函数,搭建神经网络;

[0033] S3:使用数据集对神经网络进行训练;

[0034] S4:对训练结果进行后处理,将结果转化进行位置识别;

[0035] S5:对训练好的网络进行评估;计算预测成功的比例作为准确率。

[0036] 所述评估结果的判定方法为:在每一个水平行上,预测点与真实标记点的距离值小于阈值,则判定预测成功;反之,判定预测失败。

[0037] 当阈值为10像素时,模型准确率达到95.82%。

[0038] 在一个可选实施例中,包括输入层;神经网络包括特征提取层、特征增强层和分类预测层;

[0039] 输入层,用于获取图片、图像及视频;

[0040] 特征提取层,用于图片特征提取,并在18层残差网络的基础上,对网络后三层的输出进行提取,构建网络层进行特征增强;

[0041] 特征增强层,为包括卷积、归一化及Relu的conv_bn_relu块,以对残差网络的三个输出,分别构建由多个conv_bn_relu块组成的网络,最后对三个结果进行合并,对合并后的结果再次进行多个conv_bn_relu块操作以及卷积操作,形成特征增强网络层,用于图片特征增强处理;

[0042] 分类预测层,包括全连接层,用于生成预测图。

[0043] 需要说明的是,神经网络训练过程中,在18层ResNet网络基础上,对网络后三层的输出进行提取,构建网络层进行特征增强。对于特征增强层,构建了一个由卷积,归一化,Relu组成的conv_bn_relu块,对于Resnet的三个输出,分别构建由多个conv_bn_relu块组成的网络,最后对三个结果进行合并,对合并后的结果再次进行多个conv_bn_relu块操作以及卷积操作。

[0044] 其中,特征增强层只在训练过程中使用,在验证测试时并不使用。

[0045] 所述特征提取层包括整体网络的损失函数,损失函数包括整体分类损失、整体预测损失、分割损失以及特征增强层损失,上述损失均采用交叉熵损失函数。

[0046] 对于图像中的直线目标体,利用神经网络将识别任务视为语义分割任务。

[0047] 需要说明的是,鉴于直线目标体的物理特点以及在图像中的位置特点,参考了相关研究更多的车道线检测算法。直线目标体检测和车道线检测有着很强的相似性,因此认为参考车道线检测算法并应用到直线目标体检测上也能取得良好的效果。

[0048] 综上,本发明中,鉴于2D图像中的直线目标体物理特点,以及其在图像中的位置特点,选择深度学习的方法,使用卷积神经网络进行识别;参考车道线检测算法,以ResNet作为特征提取网络,训练时加入用于特征增强的网络层,提高特征提取效果,进而提高识别效果。

[0049] 实施例1

[0050] 识别直线目标体应用在具体实例上,选择识别图片或视频中的高尔夫球杆。构建高尔夫数据集,将高尔夫球杆检测视为单直线目标体的识别,参考车道线检测算法,亦即将高尔夫球杆检测视为单车道线检测,并对识别结果进行了评估分析。具体包括以下步骤:

[0051] 步骤1:构建高尔夫球杆数据集,包括球杆标注,格式转换,拆分训练集测试集等。

[0052] 步骤2:搭建神经网络,以Qin,ZequanandWang,Huanyuand Li,Xi提出的车道线检测算法为基本框架,参考其中网络结构和损失函数。网络结构设计如下图1所示;

[0053] 其中,用于特征增强的网络层仅在训练过程中使用。分类损失及分割损失函数选择交叉熵损失函数。

[0054] 步骤3:使用构建的数据集对网络进行训练。

[0055] 步骤4:对训练结果进行后处理,将结果转化为最终高尔夫球杆识别位置。

[0056] 步骤5:对训练好的网络进行评估,在每一个水平行上,若预测的点与真实标记的点的距离小于阈值,则认为预测成功,计算预测成功的比例作为准确率。

[0057] 测试结果表明:图片的球杆和背景有着一定程度的重合,球杆部分地方在图片上显示有些模糊。提取的特征图表明,添加了特征增强的网络层的网络有效的提取到了球杆的位置信息。

[0058] 预测图表明,网络有效地识别到了高尔夫球杆位置。

[0059] 通过上述结果及类比分析,本模型可以很好的识别图像中的直线目标体。

[0060] 实施例2

[0061] 以检测高尔夫球杆为例,作为直线目标体识别的具体应用,如下:

[0062] 构建好数据集,并搭建网络后,进行训练,训练100轮。训练的分类误差以及特征增强层的误差分别如图2-3所示;模型在验证集上的表现如图4-5所示;其中,top3准确率定义为预测位置与真实标签距离小于三个单元格的数量与总单元格数量之比。从图4和图5中可以看出,两个评价指标都随训练过程明显提升。

[0063] 训练好的网络在测试集上进行验证;定义模型准确率为:预测的点与真实标记的点的距离小于阈值数量比上总数。取阈值为10,模型的准确率能够达到95.82%。

[0064] 本文中所描述的具体实施例仅仅是对本发明精神作举例说明。本发明所属技术领域的技术人员可以对所描述的具体实施例做各种各样的修改或补充或采用类似的方式替代,但并不会偏离本发明的精神或者超越所附权利要求书所定义的范围。

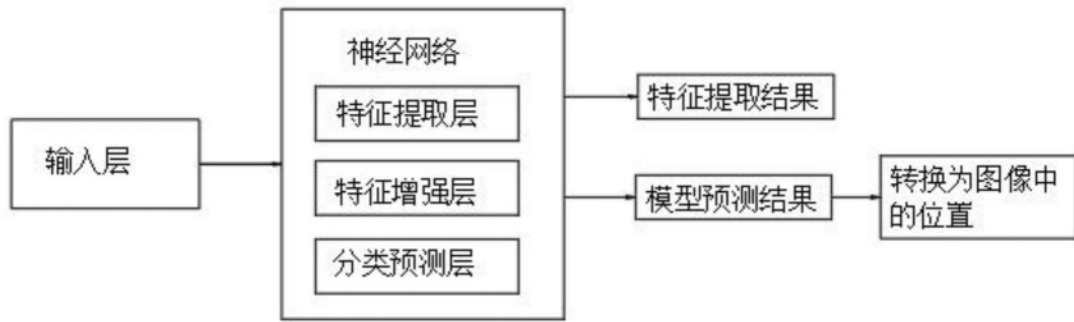


图1

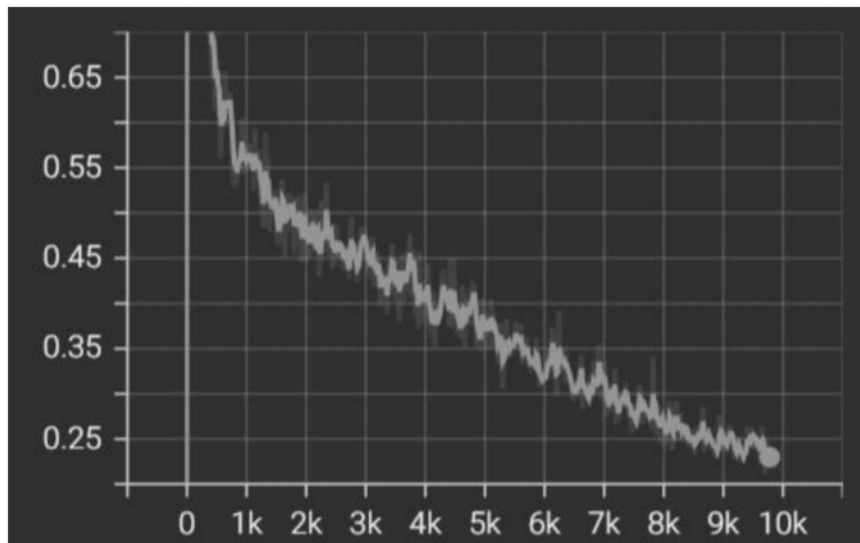


图2

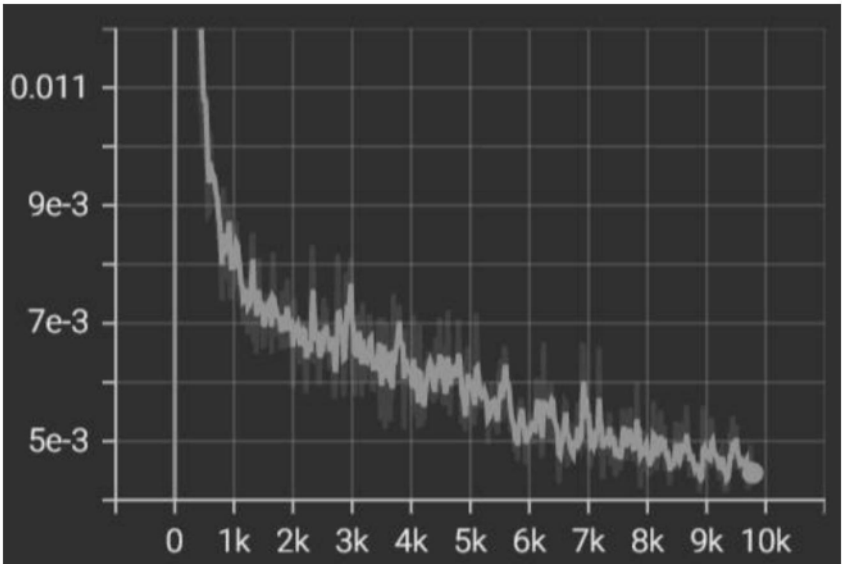


图3

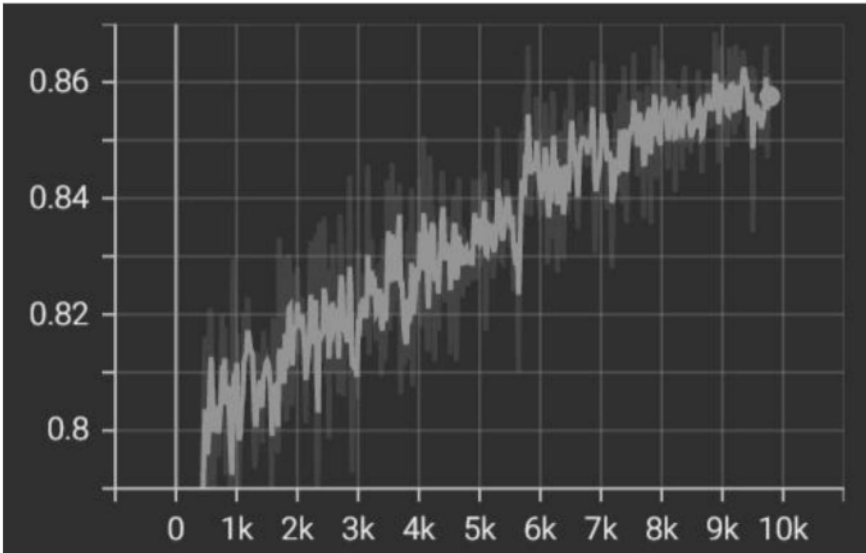


图4

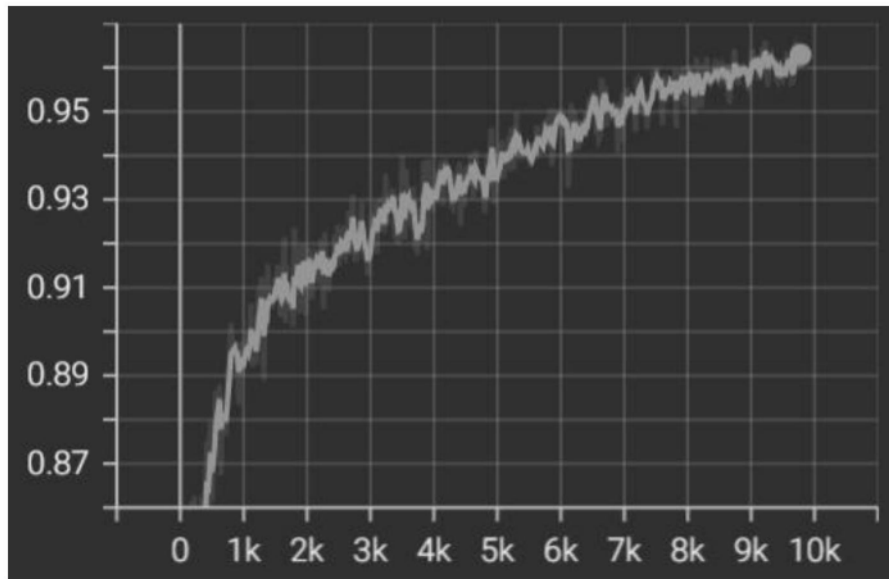


图5