



# (12) 发明专利

(10) 授权公告号 CN 111027427 B

(45) 授权公告日 2023.07.18

(21) 申请号 201911195639.9

G06N 3/0464 (2023.01)

(22) 申请日 2019.11.29

G06N 3/08 (2023.01)

(65) 同一申请的已公布的文献号  
申请公布号 CN 111027427 A

### (56) 对比文件

CN 108090423 A, 2018.05.29

CN 110378281 A, 2019.10.25

(43) 申请公布日 2020.04.17

CN 110276316 A, 2019.09.24

(73) 专利权人 大连理工大学

US 2018129865 A1, 2018.05.10

地址 116024 辽宁省大连市甘井子区凌工  
路2号

WO 2019199967 A1, 2019.10.17

廖佳伟 . 基于卷积神经网络的无人机目标  
检测算法研究 . 2019, 31-62.

(72) 发明人 庄严 汤仕豪 慕铮 闫飞

审查员 林松岭

(74) 专利代理机构 大连理工大学专利中心  
21200

专利代理师 梅洪玉 刘秋彤

(51) Int. Cl.

G06V 20/13 (2022.01)

G06V 10/774 (2022.01)

权利要求书2页 说明书4页 附图2页

### (54) 发明名称

一种用于小型无人机竞速比赛的目标门检测  
方法

### (57) 摘要

一种用于小型无人机竞速比赛的目标门检测  
方法,属于基于视觉的目标检测技术领域。通过  
对目标门拐角处关键点的检测,进而定位出目  
标门的位置。本发明利用了目标门的边缘信息,  
避免了目标门内部背景信息的影响。此外本发  
明设计的神经网络具有结构轻量化的特点,能  
够在小型无人机机载计算机上达到实时检测的  
效果。



1. 一种用于小型无人机竞速比赛的目标门检测方法,其特征在于,步骤如下:

#### 步骤一、深度学习网络的构建

深度学习网络包括特征提取模块和关键点预测模块;特征提取模块包括33个卷积层,分为16个3\*3的组卷积和17个卷积;其中,1个组卷积与1个卷积组成1个深度可分离卷积模块;剩余1个单一的卷积用于对输入图像进行滤波;深度可分离卷积模块中第一层使用3\*3的组卷积进行空间特征的提取,第二层使用1\*1的卷积将特征图恢复到相应大小;整个特征提取模块输出的特征图大小是输入图像尺寸的1/4;同时,在每组深度可分离卷积结束之后,使用通道混洗的方法,将输出特征图的各个通道重新排列,从而让各个组卷积之间学习到不同通道的特征;

关键点预测模块采用1个卷积层对网络提取的特征进行关键点的预测;该卷积层的输入与特征提取模块的输出一致,该卷积层的输出的通道数与关键点的数量一致;关键点预测模块输出一个多通道的热力图,热力图的某个通道与目标门上某个关键点相对应;找到预测图中每个通道最大值的索引,就能映射出每个关键点在图像中的位置;

#### 步骤二、训练数据集的构建

从各个角度采集待检测的目标门数据,数据需要覆盖真实场景中的各种情况;对于采集好的目标门数据,对图像中目标门的拐角处关键点进行标注;

#### 步骤三、深度学习网络的训练

深度学习网络的训练包括训练数据的处理和模型优化;训练数据的处理是对于一张图片随机进行裁剪、旋转、镜像和图片模糊,模型优化是在训练过程中同时随机加入不包含目标门的图片作为负样本训练数据,增加网络对图像背景信息干扰的适应能力,从而抑制网络的误检;

#### 步骤四、网络的推理预测

首先使用阈值对网络的预测图进行过滤,即高于阈值的网络输出值进行保留,低于阈值的网络输出值设置为0;然后对于过滤后的预测图,获取每个通道的峰值点的位置作为该通道对应的关键点的具体位置,当该通道的峰值为0时,则认为相应关键点不存在。

2. 根据权利要求1所述的一种用于小型无人机竞速比赛的目标门检测方法,其特征在于,所述步骤二中,对于每个关键点,生成一张对应的热力图;热力图的生成使用高斯核函数 $f(x)$ ,具体公式如下:

$$f(x) = \begin{cases} e^{-\frac{(dx^2+dy^2)}{2\delta^2}}, & \sqrt{dx^2 + dy^2} < r \\ 0, & \sqrt{dx^2 + dy^2} \geq r \end{cases}$$

其中dx表示标记中心到圆内某点的x方向的差值,dy表示标记中心到圆内某点的y方向的差值, $\delta$ 为归一化因子,r表示设置的高斯核半径;从上式可以得到,离标记中心越近响应值越大,超过设置的高斯核半径,响应值为0;对于目标门拐角处的关键点,将构成一幅n通道的真值热力图,n为目标门拐角处关键点的总数,某个通道对应某个关键点,用于监督网络的训练。

3. 根据权利要求1或2所述的一种用于小型无人机竞速比赛的目标门检测方法,其特征在于,所述步骤三中,训练过程使用Adam优化器,通过对梯度的一阶矩和二阶矩进行综合估

计,对网络进行优化;优化的目标函数为均方误差MSE,网络预测的通道图和真值热力图逐像素进行误差的计算,再除以总像素个数,MSE的计算方式具体如下所示:

$$MSE = \frac{1}{m} \sum_{i=1}^m (x_i - y_i)^2$$

其中 $x_i$ 是网络预测值, $y_i$ 是真值标签, $m$ 为预测的像素总数;使用标准正态分布初始化参数;每一个训练轮次完成之后,在验证集上对模型的进行评估,保留训练过程中验证精度最高的模型。

4. 根据权利要求1或2所述的一种用于小型无人机竞速比赛的目标门检测方法,其特征在于,所述步骤四中,关键点检测的评估方法: $d_i$ 为每个关键点对应真实值和预测值的欧式距离;

$$d_i = \sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2}$$

其中 $x_i$ 为真值关键点的x坐标, $y_i$ 为真值关键点的y坐标;其中 $\bar{x}_i$ 为预测关键点的x坐标, $\bar{y}_i$ 为预测关键点的y坐标;当距离 $d_i$ 小于设定阈值 $\mu$ ,认为预测值正确;

$$f(x) = \begin{cases} 1, & d_i < \mu \\ 0, & d_i \geq \mu \end{cases}$$

通过对关键点的预测正确数量的统计,能够得到模型预测的平均精度。

5. 根据权利要求3所述的一种用于小型无人机竞速比赛的目标门检测方法,其特征在于,所述步骤四中,关键点检测的评估方法: $d_i$ 为每个关键点对应真实值和预测值的欧式距离;

$$d_i = \sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2}$$

其中 $x_i$ 为真值关键点的x坐标, $y_i$ 为真值关键点的y坐标;其中 $\bar{x}_i$ 为预测关键点的x坐标, $\bar{y}_i$ 为预测关键点的y坐标;当距离 $d_i$ 小于设定阈值 $\mu$ ,认为预测值正确;

$$f(x) = \begin{cases} 1, & d_i < \mu \\ 0, & d_i \geq \mu \end{cases}$$

通过对关键点的预测正确数量的统计,能够得到模型预测的平均精度。

## 一种用于小型无人机竞速比赛的目标门检测方法

### 技术领域

[0001] 本发明属于基于视觉的目标检测技术领域,特别涉及到小型无人机竞速比赛中对目标门的实时检测方法。

### 背景技术

[0002] 随着无人机技术的发展,小型无人机竞速比赛也越来越多。通常小型无人机竞速比赛都是按照一定的顺序,通过指定的一系列目标门。在规定时间内,通过的目标门数量越多,得分越高。如果能够利用视觉传感器的感知数据,通过运行在机载计算机上的算法实时完成目标门的检测,并自动调整无人机的飞行姿态,能够大大减少通过目标门的时间,从而获得更好的比赛成绩。

[0003] 文献(Ren,Shaoqing,He,Kaiming,Girshick,Ross,et al.Faster R-CNN:Towards Real-Time Object Detection with Region Proposal Networks[J].IEEE Transactions on Pattern Analysis&Machine Intelligence,2015,39(6):1137-1149.)提出的两阶段目标检测算法,检测精度高。但是由于它将检测和识别分为两步,步骤复杂,计算量较大,在小型无人机的有限计算资源下难以满足实时检测的需求。

[0004] 文献(Redmon J,Divvala S,Girshick R,et al.You only look once:Unified,real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition.2016:779-788.)提出的单阶段目标检测算法,将图像中物体中心所在的特征用于目标的定位和识别。但是在小型无人机竞速比赛的目标门检测任务中,目标门的中心往往是背景信息,而背景信息对目标门的检测是一种干扰,因此基于上述算法的目标门检测效果不佳。

### 发明内容

[0005] 本发明针对上述现有小型无人机竞速比赛目标门检测存在的问题,提出了一种基于关键点的轻量化目标门检测方法。通过对目标门拐角处关键点的检测,进而定位出目标门的位置。本发明利用了目标门的边缘信息,避免了目标门内部背景信息的影响。此外本发明设计的神经网络具有结构轻量化的特点,能够在小型无人机机载计算机上达到实时检测的效果。

[0006] 本发明通过以下步骤进行目标门的检测:

[0007] 步骤一、深度学习网络的构建

[0008] 为了保证在小型无人机上的实时性,设计了一种轻量化深度学习网络模型,在保证精度的前提下,尽可能的优化检测的速度。该网络包括以下两个部分:特征提取模块和关键点预测模块。

[0009] 特征提取模块包括33个卷积层,分为16个3\*3的组卷积和17个卷积。其中,1个组卷积与1个卷积组成1个深度可分离卷积模块。深度学习网络中共有16个深度可分离卷积模块,此外还包括1个单一卷积,用于对输入图像进行滤波。

[0010] 深度可分离卷积模块如图1所示。模块中第一层使用3\*3的组卷积进行空间特征的提取,可大大减小计算开销。第二层使用1\*1的卷积将特征图恢复到相应大小。整个特征提取模块输出的特征图大小是输入图像尺寸的1/4。同时,在每组深度可分离卷积结束之后,使用通道混洗的方法,将输出特征图的各个通道重新排列,从而让各个组卷积之间学习到不同通道的特征,提升网络的精度。

[0011] 关键点预测模块为了保证运行效率,采用1个卷积层对网络提取的特征进行关键点的预测;该卷积层的输入与特征提取模块输出一致,该卷积层的输出的通道数与关键点的数量一致;该模块输出一个多通道的热力图,热力图的不同通道与目标门不同拐角处的关键点一一对应。找到预测图中每个通道最大值的索引,就可以映射出每个关键点在图像中的位置。具体的网络结构图如图2所示。

[0012] 步骤二、训练数据集的构建

[0013] 在训练神经网络之前,首先要采集一定量的数据用于网络的训练。先从各个角度采集待检测的目标门数据,数据需要尽量覆盖真实场景中的各种情况。对于采集好的目标门数据,对图像中目标门的拐角处关键点进行标注。

[0014] 热力图是一种密度图,使用显著的颜色差异用来显示所表达数据的分布不同。热力图可以在二维平面上直观的表现图像中某个像素与图像关键点之间的距离。如果该像素距离关键点区域越近,热力图的响应值越大;距离关键点越远,热力图的响应值越小。通过获取热力图中最大值的索引就可以获得相应的关键点位置。

[0015] 对于每个关键点,生成一张对应的热力图。热力图的生成使用高斯核函数 $f(x)$ ,具体公式如下:

$$[0016] \quad f(x) = \begin{cases} e^{-\frac{(dx^2+dy^2)}{2\delta^2}}, & \sqrt{dx^2 + dy^2} < r \\ 0, & \sqrt{dx^2 + dy^2} \geq r \end{cases}$$

[0017] 其中 $dx$ 表示标记中心到圆内某点的 $x$ 方向的差值, $dy$ 表示标记中心到圆内某点的 $y$ 方向的差值, $\delta$ 为归一化因子, $r$ 表示设置的高斯核半径。从上式可以得到,离标记中心越近响应值越大,超过设置的高斯核半径,响应值为0。对于目标门拐角处的关键点,将构成一幅 $n$ 通道的真值热力图( $n$ 为目标门拐角处关键点的总数),某个通道对应某个关键点,用于监督网络的训练。

[0018] 步骤三、深度学习网络的训练

[0019] 对网络的训练包括两部分:训练数据的处理和模型优化。在训练的过程中,对于一张图片随机进行裁剪、旋转、镜像和图片模糊等操作,这些操作能够丰富训练集样本的数量,减小网络过拟合的风险,增加网络的定位能力。同时为了克服关键点的误检问题,随机加入不包含目标门的图片作为负样本训练数据,增加网络对图像背景信息干扰的适应能力,从而抑制网络的误检。

[0020] 训练过程使用Adam优化器(Adam是一种对随机目标函数执行一阶梯度优化的算法,该算法基于适应性低阶矩估计),通过对梯度的一阶矩和二阶矩进行综合估计,对网络进行优化。优化的目标函数为MSE(Mean Squared Error,均方误差),网络预测的通道图和真值热力图逐像素进行误差的计算,再除以总像素个数,MSE的计算方式具体如下所示:

$$[0021] \quad MSE = \frac{1}{m} \sum_{i=1}^m (x_i - y_i)^2$$

[0022] 其中 $x_i$ 是网络预测值, $y_i$ 是真值标签, $m$ 为预测的像素总数。使用标准正态分布初始化参数。每一个训练轮次完成之后,在验证集上对模型的进行评估,保留训练过程中验证精度最高的模型。

[0023] 步骤四、网络的推理预测

[0024] 需要对网络输出的结果进行后处理。首先使用阈值对网络的预测图进行过滤,即高于阈值的网络输出值进行保留,低于阈值的网络输出值设置为0。然后对于过滤后的预测图,获取每个通道的峰值点的位置作为该通道对应的关键点的具体位置,如果该通道的峰值为0,则认为相应关键点不存在。

[0025] 定义了一种关键点检测的评估方法。 $d_i$ 为每个关键点对应真实值和预测值的欧式距离。

$$[0026] \quad d_i = \sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2}$$

[0027] 其中 $x_i$ 为真值关键点的x坐标, $y_i$ 为真值关键点的y坐标。其中 $\bar{x}_i$ 为预测关键点的x坐标, $\bar{y}_i$ 为预测关键点的y坐标。当距离 $d_i$ 小于设定阈值 $\mu$ ,认为预测值正确。

$$[0028] \quad f(x) = \begin{cases} 1, & di < \mu \\ 0, & di \geq \mu \end{cases}$$

[0029] 通过对关键点的预测正确数量的统计,能够得到模型预测的平均精度。

[0030] 本发明的有益效果:模型泛化能力强,在光线变化显著、光线照射不均匀等传统视觉方法检测困难的场景,有很好的鲁棒性。同时模型的结构简单,参数量小,在小型无人机载计算机上能够完成实时目标门检测任务。

## 附图说明

[0031] 图1为深度可分离卷积。

[0032] 图2为网络结构图。

## 具体实施方式

[0033] 以下结合技术方案和附图详细叙述本发明的具体实施方式。

[0034] 本方案选择大疆M210无人机作为实验平台。无人机使用单目摄像头采集图片。使用Nvidia jx2对采集来的图像进行目标门检测,定位目标门的4个关键点辅助无人机进行目标门的穿越。

[0035] 步骤一、构建深度学习网络

[0036] 本发明使用深度学习进行目标门的关键点检测和定位。网络的具体结构如表1所示。其中Conv dw表示3\*3的组卷积,卷积的组数和输入保持一致。在构建网络时,要保证输出的特征图的分辨率,尽量减小池化操作。同时为了减小计算负担,将3\*3的卷积替换为一个深度可分离卷积模块。实验中为了平衡精度和运行效率使用的输入图像尺度为256\*256,如果希望获得更高的精度,可将输入图像适当放大。

[0037] 表1、网络结构图

[0038]	卷积/步长	卷积形状	输入尺寸
	Conv /s2	3*3*3*24	256*256*3
	Conv dw/s1 Conv / s1      x 3	3*3*24*24 1*1*24*24	128*128*24 128*128*24
	Conv dw / s1	3*3*24*48	128*128*24
	Conv /s1	1*1*48*48	128*128*48
	Conv dw/s1 Conv / s1      x 7	3*3*48*48 1*1*48*48	128*128*48 128*128*48
	Conv dw / s1	3*3*48*96	128*128*48
	Conv /s1	1*1*96*96	128*128*96
	Conv dw/s1 Conv / s1      x 3	3*3*96*96 1*1*96*96	128*128*96 128*128*96
	Conv dw / s1	3*3*96*128	128*128*96
	Conv /s1	1*1*128*128	128*128*128
	Conv /s1	1*1*128*4	128*128*128

[0040] 步骤二、构建训练数据集

[0041] 首先控制飞行器在实验场地进行飞行,对球门进行各种角度的拍摄。对于现场采集拍摄的数据,使用标注工具对数据中球门的左上右下和右上左下4个点进行标注,生成相应的热力图。将数据集中80%的数据用于训练,20%用于模型精度的验证。

[0042] 步骤三、神经网络的训练

[0043] 本发明使用Nvidia 1080Ti显卡作为训练平台,使用Pytorch作为训练框架。在Pytorch框架下构建神经学习网络,实现上述网络和训练推理过程。在网络训练时,对于前5个训练轮次,学习率从0.00001逐步增加到0.0001,然后每当训练30个轮次之后,学习率缩小0.5倍。总共训练120个轮次。

[0044] 步骤四、测试网络

[0045] 本发明在网络训练完成之后,对结果进行了测试。测试的评价指标如前所述。网络在验证集上,检测精度能够达到98%,完全满足实际需求。在输出图片为256\*256的尺度下,在实验用无人机机载计算机上能够达到30帧每秒的目标门检测频率。

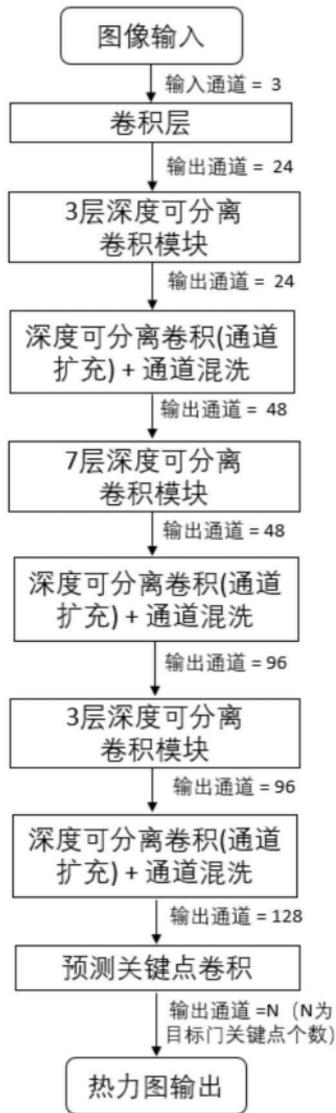


图1



图2