(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113128465 A
(43)申请公布日 2021.07.16

(21)申请号 202110508078.4

(22)申请日 2021.05.11

(71)申请人 济南大学
地址 250000 山东省济南市南辛庄西路336号

(72)发明人 李忠涛 袁朕鑫 赵帅 赵富 孙豪坤 菅国栋 李帅 姜琳琳 肖鑫 程衍泽 张玉璘 赵秀阳 孔祥玉 郭庆北 王凯

(74)专利代理机构 山东国诚精信专利代理事务所(特殊普通合伙) 37312
代理人 吴佳佳

(51)Int.Cl.
G06K 9/00(2006.01)
G06K 9/62(2006.01)
G06N 3/04(2006.01)
G06N 3/08(2006.01)

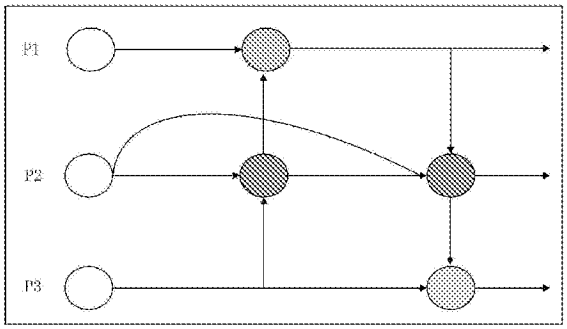
权利要求书2页 说明书5页 附图2页

(54)发明名称

一种针对工业场景基于改进YOLOv4的小目标检测方法

(57)摘要

本发明公开了一种针对工业场景基于改进YOLOv4的小目标检测方法，包括以下步骤：通过工业摄像头采集现场数据集；对采集数据集进行标注，划分训练集和测试集；通过K-Means++聚类算法，对数据集中真实目标框计算出针对本数据集的不同大小的先验框大小；对YOLOv4模型中网络进行修改，得到YOLOv4-head2网络模型；将训练集输入YOLOv4-head2模型训练，并使用验证集不断优化模型效果；在YOLOv4的特征融合层，添加与低特征层的特征融合，得到YOLOv4-head2-L网络结构；将评价较高的模型经过转换，采用TensorRT技术进行前向推理以达到提高检测速度和部署小型边缘计算设备目标。本发明具备保留与真实目标尺度接近的检测尺度，充分利用模型的网络结构，充分计算，增加低层和高层的跨层连接，更加轻量和简洁高效的优点。



CN 113128465 A

1.一种针对工业场景基于改进YOLOv4的小目标检测方法,其特征在于,包括以下步骤:

S1、通过工业摄像头采集现场数据集,采集的数据集包括在不同的天气环境光线下,丰富数据集的多样性;

S2、对采集的数据集进行标注,划分训练集和测试集,并对采集的数据集进行数据增强;

S3、通过K-Means++聚类算法,对数据集中真实目标框计算出针对本数据集的不同大小的先验框大小;

S4、对YOLOv4模型中的网络进行修改,去除其中一个的检测头,得到YOLOv4-head2网络模型;

S5、将训练集输入YOLOv4-head2模型训练,并使用验证集不断优化模型效果;

S6、在YOLOv4的特征融合层,添加与低特征层的特征融合,加强特征融合,降低随层数的增加造成的特征损失,得到YOLOv4-head2-L网络结构;

S7、将评价较高的模型经过转换,采用TensorRT技术进行前向推理以达到提高检测速度和部署小型边缘计算设备的目标。

2.根据权利要求1所述的一种针对工业场景基于改进YOLOv4的小目标检测方法,其特征在于:步骤S4中所述对模型网络进行修改,去除其中一个检测头,得到YOLOv4-L网络模型如下:

S41:根据工业生产线上目标检测的特点,图像采集摄像头的位置一般是固定的,采集的图像大小尺寸是固定的,且被识别的目标大小是在一定尺度范围内变化;

S42:YOLOv4模型中设有三个尺度的检测头,分别用于检测不同尺度大小的目标;

S43:实际目标尺度大小都是均匀尺寸的物体,因此会出现在检测时,不同尺寸的预测和分类得不到很好的训练而使得网络使用不完全;

S44:可根据实际输入图像中目标尺寸的范围,修改YOLOv4模型的检测头数量,去除冗余的检测尺度,即与预测和分类目标尺度相差较大的一支,降低模型的计算压力,降低计算量和参数量,充分利用网络结构,加快模型的检测速度;

S45:根据S3中聚类后的目标框的宽w_i和高h_i,S_1、S_2、S_3分别是YOLOV4的三个特征层的检测尺度(S_{wi}为第i个检测尺度的宽,S_{hi}为第i个检测尺度的高)。

$S_1>S_2>S_3$,$S_i=S_{wi}\cdot S_{hi}$

$$\forall w_i\in(S_{w1},S_{w2}),\quad \forall h_i\in(S_{h1},S_{h2})\ (i=1,2,\ldots,9)$$

S46:计算可推聚类出数据集中目标框的平均面积$area_{avg}$

$$area_{avg}=\frac{(\sum_{i=1}^{k}w_i\cdot h_i)}{k}$$

$$area_{avg}\in(S_{w1}\cdot S_{h1},S_{w2}\cdot S_{h2})$$

S47:由此根据S46中聚类结果的尺寸范围,对S_1、S_2、S_3不同尺度的检测头进行调整,以优化网络,防止计算资源浪费,提高模型的检测速度。

3.根据权利要求1所述的一种针对工业场景基于改进YOLOv4的小目标检测方法,其特征在于:步骤S6中所述在YOLOv4的特征融合层,添加与低特征层的特征融合特征如下:

S61:融合不同层特征有助于提高分割性能,低层特征分辨率更高,含有更多位置和细

节信息，高层特征具有更强的语义信息，分辨率较低，对细节的感知能力不强，因此低层与高层的特征融合，有助于改善模型的性能；

S62：YOLOv4网络结构中基于PANet进行实例分割，使其能够准确地保存空间信息，有助于正确定位像素点。在FPN的基础上，增加自底向上的路径增强，对不同层进行拼接，以提高预测的准确性；

S63：参考BiFPN结构，在PANet的基础上，增加低层和高层的跨层连接，与NAS-FPN相比更加轻量和简洁高效，同时更好的进行融合低层定位信息和高层的强语义信息。

4.根据权利要求1所述的一种针对工业场景基于改进YOLOv4的小目标检测方法，其特征在于：步骤S7中将评价较高的模型进过转换，使用TensorRT前向推理的过程如下：

S71：经过步骤S6得到评价较高的模型，设置batch_size和subdivisions的值为1，设置精度为INT8，将weights格式的权重文件转换为onnx格式的模型；构造推理引擎，将onnx格式权重文件转换为engine格式权重文件，engine格式权重用于模型的前向推理。

S72：将网络模型进行裁剪，对网络结构进行垂直整合，即conv、BN、Relu融合为一层；水平整合，即将输入为相同的张量和执行相同操作的融合在一起；减少concat层，将contact的输入直接送至下一步操作中，减少数据的传输吞吐。

S73：小型边缘设备部署，受限于小型设备的计算能力和内存空间有限，同时权重文件的加载速度影响着程序运行的速度，将推理引擎加载后的对象进行序列化，序列化后存至存储设备中，下次需要加载推理引擎时，只需要提取存储设备中对象反序列化后即可进行推理计算。

一种针对工业场景基于改进YOLOv4的小目标检测方法

技术领域

[0001] 本发明涉及图像识别技术领域，具体为一种针对工业场景基于改进YOLOv4的小目标检测方法。

背景技术

[0002] 在信息技术飞速发展的今天，物联网、计算机视觉技术不断应用到各个领域当中，不仅提高了生产效率。尤其是在工业领域，随着计算机视觉技术的不断更迭，在工业领域应用的方面也越来越广。工业领域的生产线上，对生产线上的产品缺陷检测之类的应用中，大部分被检测目标为小目标，图像采集设备的位置通常是固定的，计算较为复杂，效率较低，无法有效的融合低层定位信息和高层的强语义信息。对于此特点，这对模型检测小目标提出了要求。

发明内容

[0003] 本发明的目的在于提供一种针对工业场景基于改进YOLOv4的小目标检测方法，具备根据工业生产线上目标的尺寸大小和YOLOv4三个检测特征层的尺度大小，去除冗余的检测尺度，保留与真实目标尺度接近的检测尺度，充分利用模型的网络结构，充分计算，增加低层和高层的跨层连接，更加轻量和简洁高效，同时更好的进行融合低层定位信息和高层的强语义信息的优点，解决了工业领域的生产线上，对生产线上的产品缺陷检测之类的应用中，大部分被检测目标为小目标，图像采集设备的位置通常是固定的，计算较为复杂，效率较低，无法有效的融合低层定位信息和高层的强语义信息的问题。

[0004] 为实现上述目的，本发明提供如下技术方案：

[0005] 一种针对工业场景基于改进YOLOv4的小目标检测方法，包括以下步骤：

[0006] S1、通过工业摄像头采集现场数据集(包括不同类型的小目标)，采集的数据集包括在不同的天气环境光线下，丰富数据集的多样性；

[0007] S2、对采集的数据集进行标注，划分训练集和测试集，并对采集的数据集进行数据增强；

[0008] S3、通过K-Means++聚类算法，对数据集中真实目标框计算出针对本数据集的不同大小的先验框大小；

[0009] S4、对YOLOv4模型中的网络进行修改，去除其中一个的检测头，得到YOLOv4-head2网络模型；

[0010] S5、将训练集输入YOLOv4-head2模型训练，并使用验证集不断优化模型效果；

[0011] S6、在YOLOv4的特征融合层，添加与低特征层的特征融合，加强特征融合，降低随层数的增加造成的特征损失，得到YOLOv4-head2-L网络结构；

[0012] S7、将评价较高的模型经过转换，采用TensorRT技术进行前向推理以达到提高检测速度和部署小型边缘计算设备的目标。

[0013] 优选的，步骤S4中所述对模型网络进行修改，去除其中一个检测头，得到YOLOv4-L

网络模型如下：

[0014] S41：根据工业生产线上目标检测的特点，图像采集摄像头的位置一般是固定的，采集的图像大小尺寸是固定的，且被识别的目标大小是在一定尺度范围内变化；

[0015] S42：YOLOv4模型中设有三个尺度的检测头，分别用于检测不同尺度大小的目标；

[0016] S43：实际目标尺度大小都是均匀尺寸的物体，因此会出现在检测时，不同尺寸的预测和分类得不到很好的训练而使得网络使用不完全；

[0017] S44：可根据实际输入图像中目标尺寸的范围，修改YOLOv4模型的检测头数量，去除冗余的检测尺度，即与预测和分类目标尺度相差较大的一支，降低模型的计算压力，降低计算量和参数量，充分利用网络结构，加快模型的检测速度；

[0018] S45：根据S3中聚类后的目标框的宽w_i和高h_i，S_1、S_2、S_3分别是YOLOV4的三个特征层的检测尺度(S_{wi}为第i个检测尺度的宽，S_{hi}为第i个检测尺度的高)。

[0019] $S_1>S_2>S_3$，$S_i=S_{wi}\cdot S_{hi}$

[0020] $\forall w_i\in(S_{w1},S_{w2})$，$\forall h_i\in(S_{h1},S_{h2})$ (i=1,2,...,9)

[0021] S46：计算可推聚类出数据集中目标框的平均面积$area_{avg}$

[0022] $$area_{avg}=\frac{(\sum_{i=1}^{k}w_i\cdot h_i)}{k}$$

[0023] $area_{avg}\in(S_{w1}\cdot S_{h1},S_{w2}\cdot S_{h2})$

[0024] S47：由此根据S46中聚类结果的尺寸范围，对S_1、S_2、S_3不同尺度的检测头进行调整，以优化网络，防止计算资源浪费，提高模型的检测速度。

[0025] 优选的，步骤S6中所述在YOLOv4的特征融合层，添加与低特征层的特征融合特征如下：

[0026] S61：融合不同层特征有助于提高分割性能，低层特征分辨率更高，含有更多位置和细节信息，高层特征具有更强的语义信息，分辨率较低，对细节的感知能力不强，因此低层与高层的特征融合，有助于改善模型的性能；

[0027] S62：YOLOv4网络结构中基于PANet进行实例分割，使其能够准确地保存空间信息，有助于正确定位像素点。在FPN的基础上，增加自底向上的路径增强，对不同层进行拼接，以提高预测的准确性；

[0028] S63：参考BiFPN结构，在PANet的基础上，增加低层和高层的跨层连接，与NAS-FPN相比更加轻量和简洁高效，同时更好的进行融合低层定位信息和高层的强语义信息。

[0029] 优选的，步骤S7中将评价较高的模型进过转换，使用TensorRT前向推理的过程如下：

[0030] S71：经过步骤S6得到评价较高的模型，设置batch_size和subdivisions的值为1(即以单张图像输入进行预测)，设置精度为INT8，将weights格式的权重文件转换为onnx格式的模型；构造推理引擎，将onnx格式权重文件转换为engine格式权重文件，engine格式权重用于模型的前向推理。

[0031] S72：将网络模型进行裁剪，对网络结构进行垂直整合，即conv、BN、Relu融合为一层；水平整合，即将输入为相同的张量和执行相同操作的融合在一起；减少concat层，将contact的输入直接送至下一步操作中，减少数据的传输吞吐。

[0032] S73:小型边缘设备部署,受限于小型设备的计算能力和内存空间有限,同时权重文件(以下均成为推理引擎)的加载速度影响着程序运行的速度,将推理引擎加载后的对象进行序列化,序列化后存至存储设备中,下次需要加载推理引擎时,只需要提取存储设备中对象反序列化后即可进行推理计算。

[0033] 与现有技术相比,本发明的有益效果是:根据工业生产线上目标的尺寸大小和YOLOv4三个检测特征层的尺度大小,去除冗余的检测尺度,保留与真实目标尺度接近的检测尺度,充分利用模型的网络结构,充分计算;参考BiFPN结构,在PANet的基础上,增加低层和高层的跨层连接,与NAS-FPN相比更加轻量和简洁高效,同时更好的进行融合低层定位信息和高层的强语义信息;实验结果表明,在不影响检测速度的前提下,通过加强特征融合和去除冗余检测头,改进后的YOLOv4-head2-L在测试集上的AP75高达79.3%,在测试工业应用单目标检测的准确率比原YOLOv4算法提高了7.7%。

附图说明

[0034] 图1为本发明的修改后的PANet结构;

[0035] 图2为本发明的改进后的YOLOv4-L网络结构;

[0036] 图3为本发明的修改后的网络结构,实验对比。

具体实施方式

[0037] 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,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术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。

[0038] 在本发明的描述中,需要说明的是,术语“上”、“下”、“内”、“外”“前端”、“后端”、“两端”、“一端”、“另一端”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性。

[0039] 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“安装”、“设置有”、“连接”等,应做广义理解,例如“连接”,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;可以是直接相连,也可以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,可以具体情况理解上述术语在本发明中的具体含义。

[0040] 本发明提供的一种针对工业场景基于改进YOLOv4的小目标检测方法的技术方案:

[0041] 实施例1:

[0042] 一种针对工业场景基于改进YOLOv4的小目标检测方法,包括以下步骤:

[0043] S1、通过工业摄像头采集现场数据集(包括不同类型的小目标),采集的数据集包括在不同的天气环境光线下,丰富数据集的多样性;

[0044] S2、对采集的数据集进行标注,划分训练集和测试集,并对采集的数据集进行数据增强;

[0045] S3、通过K-Means++聚类算法，对数据集中真实目标框计算出针对本数据集的不同大小的先验框大小；

[0046] S4、对YOLOv4模型中的网络进行修改，去除其中一个的检测头，得到YOLOv4-head2网络模型；

[0047] S5、将训练集输入YOLOv4-head2模型训练，并使用验证集不断优化模型效果；

[0048] S6、在YOLOv4的特征融合层，添加与低特征层的特征融合，加强特征融合，降低随层数的增加造成的特征损失，得到YOLOv4-head2-L网络结构；

[0049] S7、将评价较高的模型经过转换，采用TensorRT技术进行前向推理以达到提高检测速度和部署小型边缘计算设备的目标。

[0050] 实施例2：

[0051] 步骤S4中所述对模型网络进行修改，去除其中一个检测头，得到YOLOv4-L网络模型如下：

[0052] S41：根据工业生产线上目标检测的特点，图像采集摄像头的位置一般是固定的，采集的图像大小尺寸是固定的，且被识别的目标大小是在一定尺度范围内变化；

[0053] S42：YOLOv4模型中设有三个尺度的检测头，分别用于检测不同尺度大小的目标；

[0054] S43：实际目标尺度大小都是均匀尺寸的物体，因此会出现在检测时，不同尺寸的预测和分类得不到很好的训练而使得网络使用不完全；

[0055] S44：可根据实际输入图像中目标尺寸的范围，修改YOLOv4模型的检测头数量，去除冗余的检测尺度，即与预测和分类目标尺度相差较大的一支，降低模型的计算压力，降低计算量和参数量，充分利用网络结构，加快模型的检测速度；

[0056] S45：根据S3中聚类后的目标框的宽w_i和高h_i，S_1、S_2、S_3分别是YOLOV4的三个特征层的检测尺度(S_{wi}为第i个检测尺度的宽，S_{hi}为第i个检测尺度的高)。

[0057] $S_1>S_2>S_3$，$S_i=S_{wi}\cdot S_{hi}$

[0058] $\forall w_i\in(S_{w1},S_{w2})$，$\forall h_i\in(S_{h1},S_{h2})$ (i=1,2,...,9)

[0059] S46：计算可推聚类出数据集中目标框的平均面积$area_{avg}$

[0060] $$area_{avg}=\frac{(\sum_{i=1}^{k}w_i\cdot h_i)}{k}$$

[0061] $area_{avg}\in(S_{w1}\cdot S_{h1},S_{w2}\cdot S_{h2})$

[0062] S47：由此根据S46中聚类结果的尺寸范围，对S_1、S_2、S_3不同尺度的检测头进行调整，以优化网络，防止计算资源浪费，提高模型的检测速度。

[0063] 实施例3：

[0064] 步骤S6中所述在YOLOv4的特征融合层，添加与低特征层的特征融合特征如下：

[0065] S61：融合不同层特征有助于提高分割性能，低层特征分辨率更高，含有更多位置和细节信息，高层特征具有更强的语义信息，分辨率较低，对细节的感知能力不强，因此低层与高层的特征融合，有助于改善模型的性能；

[0066] S62：YOLOv4网络结构中基于PANet进行实例分割，使其能够准确地保存空间信息，有助于正确定位像素点。在FPN的基础上，增加自底向上的路径增强，对不同层进行拼接，以提高预测的准确性；

[0067] S63:参考BiFPN结构,在PANet的基础上,增加低层和高层的跨层连接,与NAS-FPN相比更加轻量和简洁高效,同时更好的进行融合低层定位信息和高层的强语义信息。

[0068] 实施例4:

[0069] 步骤S7中将评价较高的模型进过转换,使用TensorRT前向推理的过程如下:

[0070] S71:经过步骤S6得到评价较高的模型,设置batch_size和subdivisions的值为1(即以单张图像输入进行预测),设置精度为INT8,将weights格式的权重文件转换为onnx格式的模型;构造推理引擎,将onnx格式权重文件转换为engine格式权重文件,engine格式权重用于模型的前向推理。

[0071] S72:将网络模型进行裁剪,对网络结构进行垂直整合,即conv、BN、Relu融合为一层;水平整合,即将输入为相同的张量和执行相同操作的融合在一起;减少concat层,将contact的输入直接送至下一步操作中,减少数据的传输吞吐。

[0072] S73:小型边缘设备部署,受限于小型设备的计算能力和内存空间有限,同时权重文件(以下均成为推理引擎)的加载速度影响着程序运行的速度,将推理引擎加载后的对象进行序列化,序列化后存至存储设备中,下次需要加载推理引擎时,只需要提取存储设备中对象反序列化后即可进行推理计算。

[0073] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

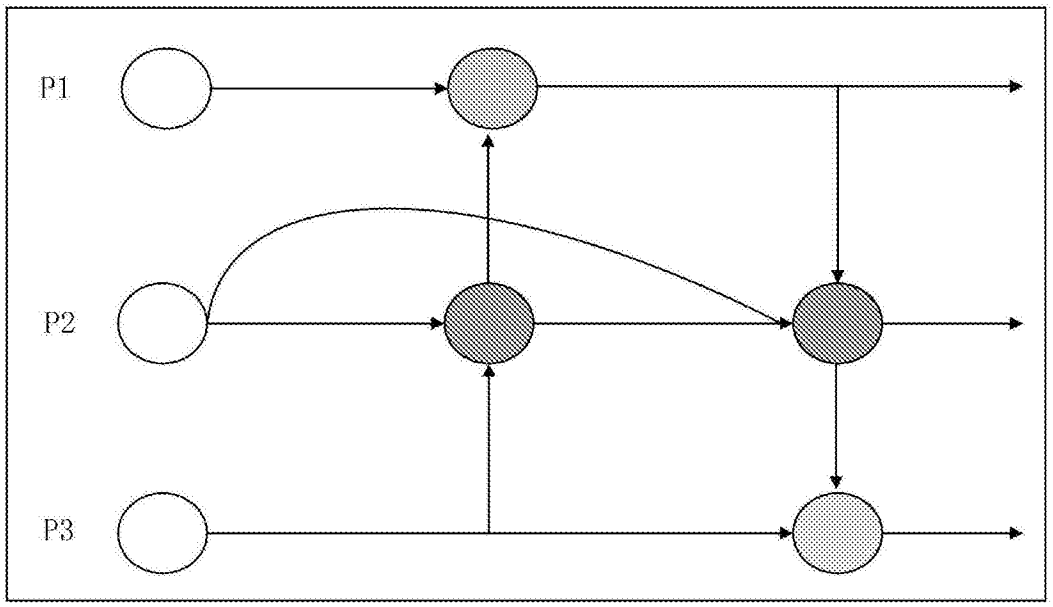

P1
P2
P3

图1

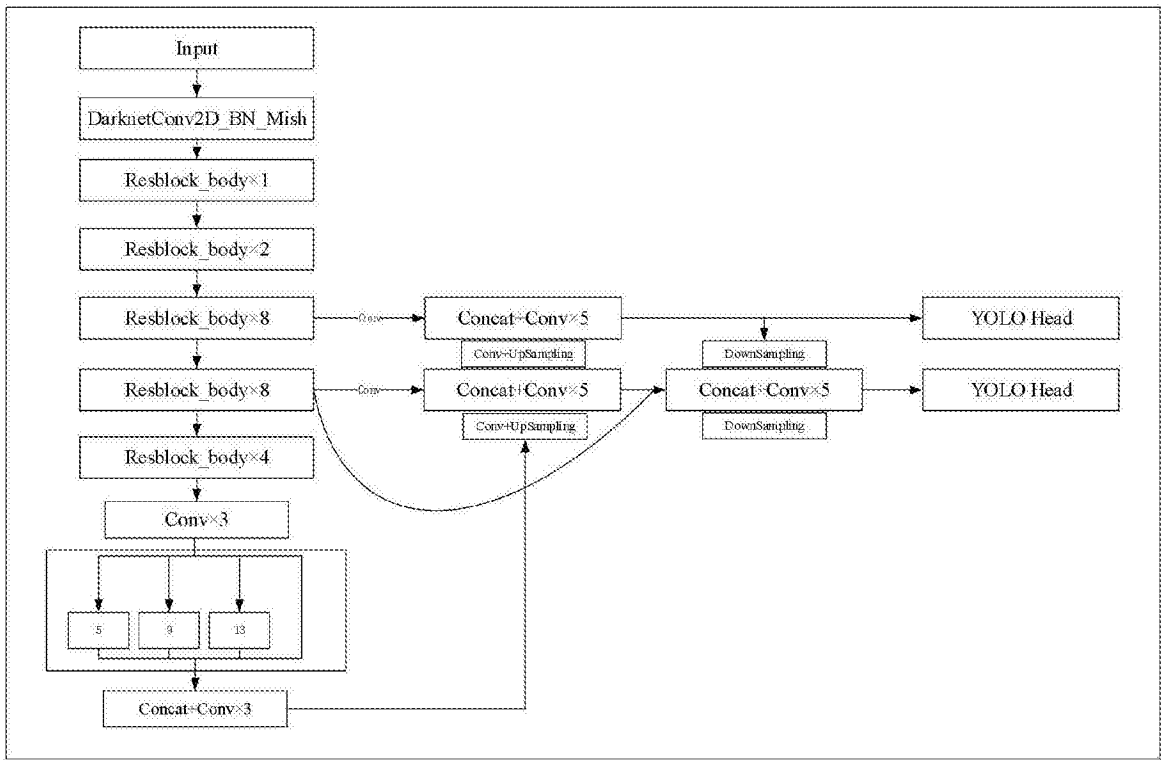

Input
DarknetConv2D_BN_Mish
Resblock_body×1
Resblock_body×2
Resblock_body×8
Resblock_body×8
Resblock_body×4
Conv×3
5
9
13
Concat+Conv×3
Conv
Concat+Conv×5
Conv+UpSampling
Conv
Concat+Conv×5
Conv+UpSampling
DownSampling
Concat+Conv×5
DownSampling
YOLO Head
YOLO Head

图2

Model	Input Size(px×px)	AP75	FPS
YOLOv4	416×416	69.5	47
YOLOv4-head2	416×416	70.3	49
YOLOv4-head2-L	416×416	**73.8**	**45**
YOLOv4	512×512	71.6	35
YOLOv4-head2	512×512	74.7	40
YOLOv4-head2-L	512×512	**79.3**	**35**

图3