



(12) 发明专利申请

(10) 申请公布号 CN 115482529 A

(43) 申请公布日 2022. 12. 16

(21) 申请号 202211179193.2

(22) 申请日 2022.09.26

(71) 申请人 中南民族大学

地址 430000 湖北省武汉市洪山区民族大道708号、823号

申请人 湖北省农业科学院农业经济技术研究所(湖北省农业科学院农业规划设计研究所)

(72) 发明人 帖军 赵捷 郑禄 王江晴

沈祥成 罗治情 官波 莫海芳 洪博文

(74) 专利代理机构 深圳市世纪恒程知识产权代

理事务所 44287

专利代理师 何秋石

(51) Int. Cl.

G06V 20/68 (2022.01)

G06V 10/764 (2022.01)

G06V 10/774 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

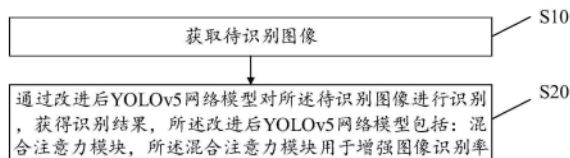
权利要求书2页 说明书10页 附图5页

(54) 发明名称

近景色水果图像识别方法、设备、存储介质及装置

(57) 摘要

本发明涉及图像处理技术领域,公开了一种近景色水果图像识别方法、设备、存储介质及装置,该方法包括:获取待识别图像,通过改进后YOLOv5网络模型对待识别图像进行识别,获得识别结果,改进后YOLOv5网络模型包括:混合注意力模块,混合注意力模块用于增强图像识别率;由于本发明改进后YOLOv5网络模型的主干网络添加了混合注意力模块,从而能够增强图像的特征信息,提高图像识别率。



1. 一种近景色水果图像识别方法,其特征在于,所述近景色水果图像识别方法包括以下步骤:

获取待识别图像;

通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

2. 如权利要求1所述的近景色水果图像识别方法,其特征在于,所述混合注意力模块包括:通道注意力SE模块和协同注意力CA模块,所述通道注意力SE模块设置在所述改进后YOLOv5网络模型的第一卷积层之后,所述协同注意力CA模块设置在所述改进后YOLOv5网络模型的第三C3模块之前,所述通道注意力SE模块用于确定通道间的信息,所述协同注意力CA模块用于确定方向感知信息和位置感知信息。

3. 如权利要求2所述的近景色水果图像识别方法,其特征在于,所述通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果的步骤,包括:

通过改进后YOLOv5网络模型的通道注意力SE模块对所述待识别图像的各通道的权重值进行调整;

通过改进后YOLOv5网络模型的协同注意力CA模块确定所述待识别图像的方向感知特征图和位置感知特征图;

根据调整后权重值、所述方向感知特征图以及所述位置感知特征图对所述待识别图像进行识别,获得识别结果。

4. 如权利要求3所述的近景色水果图像识别方法,其特征在于,所述通过改进后YOLOv5网络模型的协同注意力CA模块确定所述待识别图像的方向感知特征图和位置感知特征图的步骤,包括:

通过改进后YOLOv5网络模型的协同注意力CA模块将坐标注意力分解为一维特征编码;

基于所述一维特征编码通过所述协同注意力CA模块沿着空间维度聚合特征,获得方向感知特征图和位置感知特征图。

5. 如权利要求1至4中任一项所述的近景色水果图像识别方法,其特征在于,所述改进后YOLOv5网络模型还包括:改进后模型特征融合连接结构,所述改进后模型特征融合连接结构设置在C3模块之前,用于进行Concat特征融合操作。

6. 如权利要求1至4中任一项所述的近景色水果图像识别方法,其特征在于,所述获取待识别图像的步骤之前,还包括:

获取图像样本,并将所述图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数;

基于所述Varifocal Loss损失函数对所述初始YOLOv5网络模型进行调整,获得改进后YOLOv5网络模型。

7. 如权利要求6所述的近景色水果图像识别方法,其特征在于,所述获取图像样本,并将所述图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数的步骤之前,还包括:

获取初始图像;

通过改进后YOLOv5模型的Mosaic数据增强模块对所述初始图像进行数据增强,获得图像样本。

8. 一种近景色水果图像识别设备,其特征在於,所述近景色水果图像识别设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的近景色水果图像识别程序,所述近景色水果图像识别程序被所述处理器执行时实现如权利要求1至7中任一项所述的近景色水果图像识别方法。

9. 一种存储介质,其特征在於,所述存储介质上存储有近景色水果图像识别程序,所述近景色水果图像识别程序被处理器执行时实现如权利要求1至7中任一项所述的近景色水果图像识别方法。

10. 一种近景色水果图像识别装置,其特征在於,所述近景色水果图像识别装置包括:图像获取模块和图像识别模块;

所述图像获取模块,用于获取待识别图像;

所述图像识别模块,用于通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

## 近景色水果图像识别方法、设备、存储介质及装置

### 技术领域

[0001] 本发明涉及图像处理技术领域,尤其涉及一种近景色水果图像识别方法、设备、存储介质及装置。

### 背景技术

[0002] 随着科技进步与我国农业高质量发展,农业已由自动化转为人工智能化,传统的人工采摘方式已经无法满足农民对柑橘一类果实的采摘需求。如今,我国正在开展研究智能化的果实采摘机器人来满足实际的生产需要。研究智能果采摘机器人,能够有效节约劳动力成本,促进农业的智能化发展。计算机视觉技术是实现智能果实采摘机器人操作能力的关键技术。

[0003] 但是,在实际应用中,由于近景色水果果实与自然背景颜色相似,识别难度高,识别准确率低。例如,绿色柑橘在进行识别时,由于与叶片颜色较为相似,从而导致识别准确率低。

### 发明内容

[0004] 本发明的主要目的在于提供一种近景色水果图像识别方法、设备、存储介质及装置,旨在解决现有技术由于近景色水果果实与自然背景颜色相似,识别难度高,识别准确率低的技术问题。

[0005] 为实现上述目的,本发明提供一种近景色水果图像识别方法,所述近景色水果图像识别方法包括以下步骤:

[0006] 获取待识别图像;

[0007] 通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

[0008] 可选地,所述混合注意力模块包括:通道注意力SE模块和协同注意力CA模块,所述通道注意力SE模块设置在所述改进后YOLOv5网络模型的第一卷积层之后,所述协同注意力CA模块设置在所述改进后YOLOv5网络模型的第三C3模块之前,所述通道注意力SE模块用于确定通道间的信息,所述协同注意力CA模块用于确定方向感知信息和位置感知信息。

[0009] 可选地,所述通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果的步骤,包括:

[0010] 通过改进后YOLOv5网络模型的通道注意力SE模块对所述待识别图像的各通道的权重值进行调整;

[0011] 通过改进后YOLOv5网络模型的协同注意力CA模块确定所述待识别图像的方向感知特征图和位置感知特征图;

[0012] 根据调整后权重值、所述方向感知特征图以及所述位置感知特征图对所述待识别图像进行识别,获得识别结果。

[0013] 可选地,所述通过改进后YOLOv5网络模型的协同注意力CA模块确定所述待识别图

像的方向感知特征图和位置感知特征图的步骤,包括:

[0014] 通过改进后YOLOv5网络模型的协同注意力CA模块将坐标注意力分解为一维特征编码;

[0015] 基于所述一维特征编码通过所述协同注意力CA模块沿着空间维度聚合特征,获得方向感知特征图和位置感知特征图。

[0016] 可选地,所述改进后YOLOv5网络模型还包括:改进后模型特征融合连接结构,所述改进后模型特征融合连接结构设置在C3模块之前,用于进行Concat特征融合操作。

[0017] 可选地,所述获取待识别图像的步骤之前,还包括:

[0018] 获取图像样本,并将所述图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数;

[0019] 基于所述Varifocal Loss损失函数对所述初始YOLOv5网络模型进行调整,获得改进后YOLOv5网络模型。

[0020] 可选地,所述获取图像样本,并将所述图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数的步骤之前,还包括:

[0021] 获取初始图像;

[0022] 通过改进后YOLOv5模型的Mosaic数据增强模块对所述初始图像进行数据增强,获得图像样本。

[0023] 此外,为实现上述目的,本发明还提出一种近景色水果图像识别设备,所述近景色水果图像识别设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的近景色水果图像识别程序,所述近景色水果图像识别程序配置为实现如上文所述的近景色水果图像识别方法。

[0024] 此外,为实现上述目的,本发明还提出一种存储介质,所述存储介质上存储有近景色水果图像识别程序,所述近景色水果图像识别程序被处理器执行时实现如上文所述的近景色水果图像识别方法。

[0025] 此外,为实现上述目的,本发明还提出一种近景色水果图像识别装置,所述近景色水果图像识别装置包括:图像获取模块和图像识别模块;

[0026] 所述图像获取模块,用于获取待识别图像;

[0027] 所述图像识别模块,用于通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

[0028] 在本发明中,公开了获取待识别图像,通过改进后YOLOv5网络模型对待识别图像进行识别,获得识别结果,改进后YOLOv5网络模型包括:混合注意力模块,混合注意力模块用于增强图像识别率;由于本发明改进后YOLOv5网络模型的主干网络添加了混合注意力模块,从而能够增强图像的特征信息,提高图像识别率。

## 附图说明

[0029] 图1是本发明实施例方案涉及的硬件运行环境的近景色水果图像识别设备的结构示意图;

[0030] 图2为本发明近景色水果图像识别方法第一实施例的流程示意图;

- [0031] 图3为现有YOLOv5网络模型示意图；
- [0032] 图4为本发明近景色水果图像识别方法一实施例的改进后YOLOv5网络模型示意图；
- [0033] 图5为本发明近景色水果图像识别方法第二实施例的流程示意图；
- [0034] 图6为本发明近景色水果图像识别方法第三实施例的流程示意图；
- [0035] 图7为本发明近景色水果图像识别装置第一实施例的结构框图。
- [0036] 本发明目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

### 具体实施方式

- [0037] 应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。
- [0038] 参照图1，图1为本发明实施例方案涉及的硬件运行环境的近景色水果图像识别设备结构示意图。
- [0039] 如图1所示，该近景色水果图像识别设备可以包括：处理器1001，例如中央处理器(Central Processing Unit, CPU)，通信总线1002、用户接口1003，网络接口1004，存储器1005。其中，通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)，可选用户接口1003还可以包括标准的有线接口、无线接口，对于用户接口1003的有线接口在本发明中可为USB接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如无线保真(Wireless-Fidelity, Wi-Fi)接口)。存储器1005可以是高速的随机存取存储器(Random Access Memory, RAM)，也可以是稳定的存储器(Non-volatile Memory, NVM)，例如磁盘存储器。存储器1005可选的还可以是独立于前述处理器1001的存储装置。
- [0040] 本领域技术人员可以理解，图1中示出的结构并不构成对近景色水果图像识别设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。
- [0041] 如图1所示，认定为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及近景色水果图像识别程序。
- [0042] 在图1所示的近景色水果图像识别设备中，网络接口1004主要用于连接后台服务器，与所述后台服务器进行数据通信；用户接口1003主要用于连接用户设备；所述近景色水果图像识别设备通过处理器1001调用存储器1005中存储的近景色水果图像识别程序，并执行本发明实施例提供的近景色水果图像识别方法。
- [0043] 基于上述硬件结构，提出本发明近景色水果图像识别方法的实施例。
- [0044] 参照图2，图2为本发明近景色水果图像识别方法第一实施例的流程示意图，提出本发明近景色水果图像识别方法第一实施例。
- [0045] 步骤S10：获取待识别图像。
- [0046] 应当理解的是，本实施例的执行主体可以是具有数据处理、网络通信以及程序运行功能的近景色水果图像识别设备，例如，电脑等，或者是其他能够实现相同或相似功能的电子设备，本实施例对此不加限制。
- [0047] 需要说明的是，待识别图像可以由用户预先输入，例如，待识别图像可以是用户预先输入的绿色柑橘图像。
- [0048] 可以理解的是，为了提高处理效率，本实施例中，可以先对初始图像进行预处理，

以获得待识别图像。其中,预处理可以是去噪等处理,以提高图像质量。

[0049] 步骤S20:通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

[0050] 需要说明的是,混合注意力模块可以是通道注意力SE模块和/或协同注意力CA模块,本实施例对此不加以限制。

[0051] 识别结果可以包括待识别图像中各部分的定位和类别,例如,待识别图像为绿色柑橘图像时,识别结果为绿色柑橘的位置。

[0052] 为了便于理解,参考图3进行说明,但并不对本方案进行限定。图3为现有YOLOv5网络模型示意图,图中,YOLOv5网络的输入端存在Mosaic数据增强的策略,保证图片样本的均衡性。YOLOv5模型的主干网络主要分为Focus模块和C3模块,Focus模块将图片进入主干网络之前,对图片进行切片操作,将图像相邻地4个位置进行堆叠,把高分辨率的特征图拆分成多个低分辨率的特征图,实现下采样的同时减少计算量,并提升网络速度。C3模块是在BottleneckCSP的基础上改进而来的,它的瓶颈结构比BottleneckCSP少一个卷积层,使得模型的体积变少,concat后的卷积模块中的激活函数也换成SiLU,检测性能不断增强。Neck中采用了FPN与PAN结合的结构,其主要思想来源于PANet网络结构。FPN层自顶向下传达强语义特征,再与自底向上的特征金字塔结合,将低层特征与高层特征进行连接,增强模型的语义信息。Prediction采用CIoU\_Loss作为边界框损失函数,网络一共输出三个尺寸的特征图来进行检测,用于检测不同尺寸的目标对象,最后检测生成的图像通过预测的边界框和类别进行标记。

[0053] 为了便于理解,参考图4进行说明,但并不对本方案进行限定。图4为改进后YOLOv5网络模型示意图,图中,在现有YOLOv5网络模型的主干网络中添加混合注意力机制,即将通道注意力SE模块和协同注意力CA模块结合放入到主干网络中,不仅能捕获方向感知和位置感知信息,也能进一步加强模型对通道信息的捕获,让模型更好地提取、定位柑橘图片的特征信息,提高自然环境下柑橘识别率。

[0054] 在第一实施例中,公开了获取待识别图像,通过改进后YOLOv5网络模型对待识别图像进行识别,获得识别结果,改进后YOLOv5网络模型包括:混合注意力模块,混合注意力模块用于增强图像识别率;由于本实施例改进后YOLOv5网络模型的主干网络添加了混合注意力模块,从而能够增强图像的特征信息,提高图像识别率。

[0055] 参照图5,图5为本发明近景色水果图像识别方法第二实施例的流程示意图,基于上述图2所示的第一实施例,提出本发明近景色水果图像识别方法的第二实施例。

[0056] 在第二实施例中,所述混合注意力模块包括:通道注意力SE模块和协同注意力CA模块,所述通道注意力SE模块设置在所述改进后YOLOv5网络模型的第一卷积层之后,所述协同注意力CA模块设置在所述改进后YOLOv5网络模型的第三C3模块之前,所述通道注意力SE模块用于确定通道间的信息,所述协同注意力CA模块用于确定方向感知信息和位置感知信息。

[0057] 为了便于理解,参考图4进行说明,但并不对本方案进行限定。图4为改进后YOLOv5网络模型示意图,图中,在现有YOLOv5网络模型的主干网络中添加混合注意力机制,即在模型主干网络中的第2层嵌入SE注意力,模型主干网络的第11层嵌入CA注意力,两种注意力机

制结合一起使用,不仅能捕获方向感知和位置感知信息,也能进一步加强模型对通道信息的捕获,让模型更好地提取、定位柑橘图片的特征信息,提高自然环境下柑橘识别率。

[0058] 在第二实施例中,所述改进后YOLOv5网络模型还包括:改进后模型特征融合连接结构,所述改进后模型特征融合连接结构设置在C3模块之前,用于进行Concat特征融合操作。

[0059] 应当理解的是,卷积神经网络是专门用来处理图像数据这种类似网络结构数据的神经网络,通过卷积操作,能够更好地提取图片地特征信息。YOLOv5模型的C3模块比BottleneckCSP模块少一个卷积层,防止在卷积过程中产生的梯度爆炸和网络退化的问题。C3模块存在多次卷积操作,YOLOv5模型使用大量C3结构,模型自身也通过其他卷积操作对特征图进行操作。由于卷积操作包含大量的参数,柑橘图片与背景颜色相近,且在柑橘数量较多时,柑橘体积较小,特征提取中通过大量的卷积之后容易出现柑橘对象的特征信息丢失的问题。因此,为了克服上述缺陷,本实施例中改进后YOLOv5网络模型还包括:改进后模型特征融合连接结构,改进后模型特征融合连接结构设置在C3模块之前,用于进行Concat特征融合操作。

[0060] 为了便于理解,参考图4进行说明,但并不对本方案进行限定。图4为改进后YOLOv5网络模型示意图,YOLOv5模型采用了Concat方法将上层特征分支与下层特征分支进行特征融合,图中,为减少下层特征分支的卷积次数和Concat特征融合时的参数量,同时提高网络模型识别柑橘的性能,本文将下层特征分支改在C3模块之前,再与另一条上层特征分支进行Concat特征融合,以减少在Concat连接操作时的参数量。

[0061] 在第二实施例中,所述步骤S20,包括:

[0062] 步骤S201:通过改进后YOLOv5网络模型的通道注意力SE模块对所述待识别图像的各通道的权重值进行调整。

[0063] 应当理解的是,柑橘图片与自然环境的背景色相近,其通道特征信息有限,在柑橘数量较多的情况下,其图像的像素值也有限,所以在识别的时候需要增加通道间特征信息进行识别。

[0064] 可以理解的是,通道注意力SE模块关注通道间的信息,可以在深层次网络中优化学习特定类别的特征信息。首先,通道注意力SE模块通过压缩(Squeeze)操作,给定输入X,第C个通道的压缩操作可用如下公示表示:

$$[0065] \quad z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j)$$

[0066] 式中, $Z_c$ 是与第c个通道相关连的输出,将经过卷积操作之后的特征图进行全局平均池化操作,得到当前特征图的全局压缩特征量。

[0067] 其次,再经过激发(Excitation)操作,可用如下公示表示:

$$[0068] \quad \hat{X} = X \cdot \sigma(\hat{z})$$

[0069] 式中, $\sigma$ 表示sigmoid激活函数, $\hat{z}$ 是转换函数生成后的结果。

[0070] 转换函数生成后的结果可用下式表示:

$$[0071] \quad \hat{z} = T_2(\text{ReLU}(T_1(z)))$$

[0072] 式中, $T_1$ 和 $T_2$ 是两个线性变换来捕获每个通道的重要性,最后通过缩放操作对通



道间的特征进行加权,将通道间的权重值与原始特征图的特征相乘,得到调整后权重值。

[0073] 步骤S202:通过改进后YOLOv5网络模型的协同注意力CA模块确定所述待识别图像的方向感知特征图和位置感知特征图。

[0074] 应当理解的是,在实际应用中,识别时还需要关注位置感知信息,这有助于模型更精准地定位和识别柑橘对象。

[0075] 可以理解的是,协同注意力CA模块不仅关注了通道间的信息,同时还关注方向感知和位置感知信息。协同注意力CA模块将坐标注意力分解为两个一维特征编码,分别沿两个空间方向聚合特征,一个方向捕获远程依赖关系,另一个方向保留精准的位置信息。最终分别形成方向感知和位置感知的特征图互补地应用于输入特征图。

[0076] 步骤S203:根据调整后权重值、所述方向感知特征图以及所述位置感知特征图对所述待识别图像进行识别,获得识别结果。

[0077] 应当理解的是,在获得调整后权重值、方向感知特征图以及位置感知特征图后,可以根据调整后权重值、所述方向感知特征图以及所述位置感知特征图进行后续处理,以对待识别图像进行识别,获得识别结果。

[0078] 进一步地,所述步骤S204,包括:

[0079] 通过改进后YOLOv5网络模型的协同注意力CA模块将坐标注意力分解为一维特征编码;

[0080] 基于所述一维特征编码通过所述协同注意力CA模块沿着空间维度聚合特征,获得方向感知特征图和位置感知特征图。

[0081] 具体来看,协同注意力CA模块为了实现坐标信息嵌入,首先通过全局平均池化分别对水平X方向和垂直Y方向的每个通道进行编码,得到两个特征图。然后用Concat连接这两个特征图,然后使用一个共享的 $1 \times 1$ 卷积变换函数F1进行操作:

[0082]  $f = \delta (F_1 ([z^h, z^w]))$

[0083] 式中, $\delta$ 表示非线性激活函数, $[\cdot, \cdot]$ 表示沿空间维度的连接操作, $f \in \mathbb{R}^{C/r \times (H+W)}$ 表示水平和垂直方向对空间信息编码的中间特征映射, $r$ 表示缩放系数。然后沿着空间维度,将 $f$ 分成两个独立的张量 $f^h \in \mathbb{R}^{C/r \times H}$ 和 $f^w \in \mathbb{R}^{C/r \times W}$ 。为了使张量的通道数相同,再利用两个 $1 \times 1$ 卷积变换 $F_h$ 和 $F_w$ 分别对 $f^h$ 和 $f^w$ 进行变换,得到如下公式:

[0084]  $g^h = \sigma (F_h (f^h))$

[0085]  $g^w = \sigma (F_w (f^w))$

[0086]  $\sigma$ 表示sigmoid激活函数,这里也是用了缩放系数 $r$ 来减少 $f$ 的通道数,然后分别将扩展的 $g^h$ 和 $g^w$ 作为注意力的权重。最终,注意力模块输出的表达式可以如下所示:

[0087]  $y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j)$

[0088] 在第二实施例中,公开了通过改进后YOLOv5网络模型的通道注意力SE模块对待识别图像的各通道的权重值进行调整,通过改进后YOLOv5模型的Mosaic数据增强模块对初始图像进行数据增强,获得图像样本,根据调整后权重值、方向感知特征图以及位置感知特征图对待识别图像进行识别,获得识别结果;由于本实施例还通过通道注意力SE模块关注通道间的信息,并通过协同注意力CA模块关注方向感知和位置感知信息,从而能够进一步提高识别结果的准确性。

[0089] 参照图6,图6为本发明近景色水果图像识别方法第三实施例的流程示意图,基于

上述图2所示的第一实施例,提出本发明近景色水果图像识别方法的第三实施例。

[0090] 在第三实施例中,所述步骤S10之前,还包括:

[0091] 步骤S01:获取图像样本,并将所述图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数。

[0092] 应当理解的是,在实际应用中,绿色柑橘生长形态各异,容易受到树叶、树枝等背景物体的遮挡或自身重叠等因素影响,导致识别精度低。因此,为了克服上述缺陷,本实施例中,在对YOLOv5网络模型进行训练时,还改进了YOLOv5网络模型的损失函数,从而加强了柑橘重叠遮挡情形下的识别准确率和鲁棒性。

[0093] 需要说明的是,初始YOLOv5网络模型可以是未经训练的改进后YOLOv5网络模型。

[0094] 可以理解的是,图像识别的主要任务就是定位和分类,将图片的目标定位标注并给出相应类别。损失函数就是在神经网络训练时,用网络的预测值与实际标签信息进行比较,计算两者之间的差距。此时,网络模型不断反向迭代更新,参数不断调整更新,降低损失函数,让网络最终预测的越来越准确。因此,目标检测任务的损失函数可以分为两类,一类是分类损失函数,另一类是回归损失函数。

[0095] 现有YOLOv5网络模型分类损失函数采用的是Focal Loss损失函数,其定义如下式所示:

$$[0096] \quad FL(p, y) = \begin{cases} -\alpha(1-p)^\gamma \log(p) & \text{if } y=1 \\ -(1-\alpha)p^\gamma \log(1-p) & \text{otherwise,} \end{cases}$$

[0097] 式中, $y \in \{0, 1\}$ ,表示真实类, $p \in [0, 1]$ 表示前景类的预测概率, $\alpha$ 表示用来平衡正负样本的权重, $(1-p)^\gamma$ 和 $p^\gamma$ 用来调整每个样本的权重,Focal Loss对于正负样本的处理是相同的。

[0098] 为了提高柑橘在自然场景下的识别率,本文引入Varifocal Loss损失函数替换YOLOv5原网络模型的Focal Loss损失函数。Varifocal Loss损失函数是在Focal Loss损失函数基础上进行改进的分类损失函数,其定义如下式所示:

$$[0099] \quad FL(p, y) = \begin{cases} -q(q \log(p) + (1-q) \log(1-p)) & q > 0 \\ -\alpha p^\gamma \log(1-p) & q = 0, \end{cases}$$

[0100] 式中, $p$ 是预测的分类得分, $q$ 是目标IoU的得分, $\alpha$ 表示用来平衡正负样本的权重, $p^\gamma$ 表示缩放系数。与Focal Loss不同的是,增加系数 $q$ 来对正样本加权,如果正样本的真实IoU得分较高,则损失函数的权重会更大一些,这样可以使训练时聚焦在那些质量高的样本上。由于正样本的数量较少,为了充分利用正样本的监督信息,缩放系数 $p^\gamma$ 只会对负样本进行调节。为了平衡整体的正负样本数量,使用 $\alpha$ 对负样本进行加权。

[0101] 步骤S02:基于所述Varifocal Loss损失函数对所述初始YOLOv5网络模型进行调整,获得改进后YOLOv5网络模型。

[0102] 应当理解的是,基于Varifocal Loss损失函数对初始YOLOv5网络模型进行调整可以是基于Varifocal Loss损失函数对初始YOLOv5网络模型不断反向迭代更新,参数不断调整更新,降低损失函数,让网络最终预测的越来越准确。

[0103] 在第三实施例中,公开了获取图像样本,并将图像样本输入初始YOLOv5网络模型进行训练,获得Varifocal Loss损失函数,基于Varifocal Loss损失函数对初始YOLOv5网

络模型进行调整,获得改进后YOLOv5网络模型;由于本实施例在对YOLOv5网络模型进行训练时,还改进了YOLOv5网络模型的损失函数,从而加强了柑橘重叠遮挡情形下的识别准确率和鲁棒性。

[0104] 进一步地,为了进一步增强对小目标柑橘或柑橘遮挡情况下的训练,所述步骤S01之前,还包括:

[0105] 获取初始图像;

[0106] 通过改进后YOLOv5模型的Mosaic数据增强模块对所述初始图像进行数据增强,获得图像样本。

[0107] 应当理解的是,为了改善在某些分类场景下柑橘图片数量不够的问题,本实施例还通过改进后YOLOv5模型的Mosaic数据增强模块对初始柑橘图像进行数据增强。

[0108] 需要说明的是,初始图像可以是用户输入的柑橘图像。

[0109] 可以理解的是,YOLOv5模型存在Mosaic数据增强模块,针对每个场景下柑橘图片数量不均衡的情况,拟用Mosaic数据增强模块,以改善在某些分类场景下柑橘图片数量不够的问题,进一步增强对小目标柑橘或柑橘遮挡情况下的训练。将四张图片进行随机裁剪,再拼接到一张图上作为训练。使用Mosaic增强训练时,由于可以直接计算4张图片的数据,使得不需要增加最小批次(Mini-batch)的大小,减少了GPU的使用率,一个GPU就可以达到比较好的效果,让网络的鲁棒性更好。因此,使用该方法后的图像有利于在训练过程中更好地拟合训练集中的图像,Mosaic数据增强模块是一种仅需较小的代价就能大概率提升模型性能的训练策略。

[0110] 在第三实施例中,公开了获取初始图像,通过改进后YOLOv5模型的Mosaic数据增强模块对初始图像进行数据增强,获得待识别图像;由于本实施例还通过改进后YOLOv5模型的Mosaic数据增强模块对初始柑橘图像进行数据增强,从而能够在训练过程中更好地拟合训练集中的图像。

[0111] 此外,本发明实施例还提出一种存储介质,所述存储介质上存储有近景色水果图像识别程序,所述近景色水果图像识别程序被处理器执行时实现如上文所述的近景色水果图像识别方法。

[0112] 此外,参照图7,本发明实施例还提出一种近景色水果图像识别装置,所述近景色水果图像识别装置包括:图像获取模块10和图像识别模块20;

[0113] 所述图像获取模块10,用于获取待识别图像。

[0114] 需要说明的是,待识别图像可以由用户预先输入,例如,待识别图像可以是用户预先输入的绿色柑橘图像。

[0115] 可以理解的是,为了提高处理效率,本实施例中,可以先对初始图像进行预处理,以获得待识别图像。其中,预处理可以是去噪等处理,以提高图像质量。

[0116] 所述图像识别模块20,用于通过改进后YOLOv5网络模型对所述待识别图像进行识别,获得识别结果,所述改进后YOLOv5网络模型包括:混合注意力模块,所述混合注意力模块用于增强图像识别率。

[0117] 需要说明的是,混合注意力模块可以是通道注意力SE模块和/或协同注意力CA模块,本实施例对此不加以限制。

[0118] 识别结果可以包括待识别图像中各部分的定位和类别,例如,待识别图像为绿色

柑橘图像时,识别结果为绿色柑橘的位置。

[0119] 为了便于理解,参考图3进行说明,但并不对本方案进行限定。图3为现有YOLOv5网络模型示意图,图中,YOLOv5网络的输入端存在Mosaic数据增强的策略,保证图片样本的均衡性。YOLOv5模型的主干网络主要分为Focus模块和C3模块,Focus模块将图片进入主干网络之前,对图片进行切片操作,将图像相邻地4个位置进行堆叠,把高分辨率的特征图拆分成多个低分辨率的特征图,实现下采样的同时减少计算量,并提升网络速度。C3模块是在BottleneckCSP的基础上改进而来的,它的瓶颈结构比BottleneckCSP少一个卷积层,使得模型的体积变少,concat后的卷积模块中的激活函数也换成SiLU,检测性能不断增强。Neck中采用了FPN与PAN结合的结构,其主要思想来源于PANet网络结构。FPN层自顶向下传达强语义特征,再与自底向上的特征金字塔结合,将低层特征与高层特征进行连接,增强模型的语义信息。Prediction采用CIoU\_Loss作为边界框损失函数,网络一共输出三个尺寸的特征图来进行检测,用于检测不同尺寸的目标对象,最后检测生成的图像通过预测的边界框和类别进行标记。

[0120] 为了便于理解,参考图4进行说明,但并不对本方案进行限定。图4为改进后YOLOv5网络模型示意图,图中,在现有YOLOv5网络模型的主干网络中添加混合注意力机制,即将通道注意力SE模块和协同注意力CA模块结合放入到主干网络中,不仅能捕获方向感知和位置感知信息,也能进一步加强模型对通道信息的捕获,让模型更好地提取、定位柑橘图片的特征信息,提高自然环境下柑橘识别率。

[0121] 在本实施例中,公开了获取待识别图像,通过改进后YOLOv5网络模型对待识别图像进行识别,获得识别结果,改进后YOLOv5网络模型包括:混合注意力模块,混合注意力模块用于增强图像识别率;由于本实施例改进后YOLOv5网络模型的主干网络添加了混合注意力模块,从而能够增强图像的特征信息,提高图像识别率。

[0122] 本发明所述近景色水果图像识别装置的其他实施例或具体实现方式可参照上述各方法实施例,此处不再赘述。

[0123] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者系统不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。

[0124] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0125] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如只读存储器镜像(Read Only Memory image,ROM)/随机存取存储器(Random Access Memory,RAM)、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本发明各个实施例所述的方法。

[0126] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技

---

术领域,均同理包括在本发明的专利保护范围内。

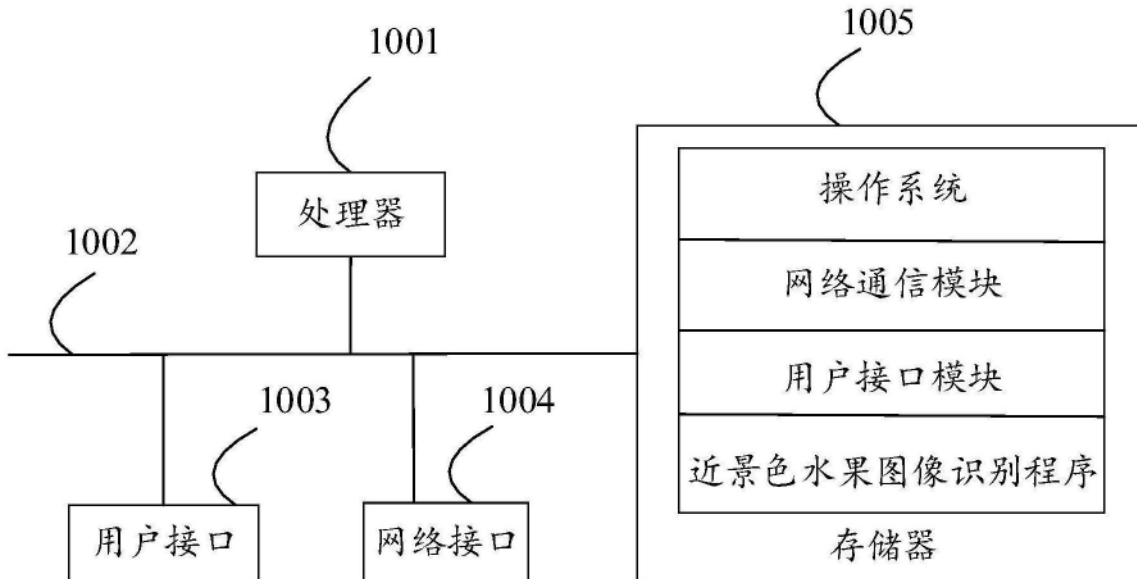


图1

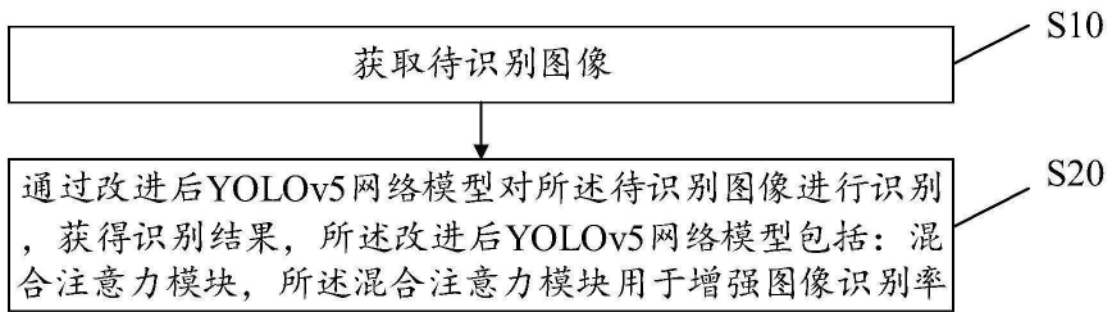


图2

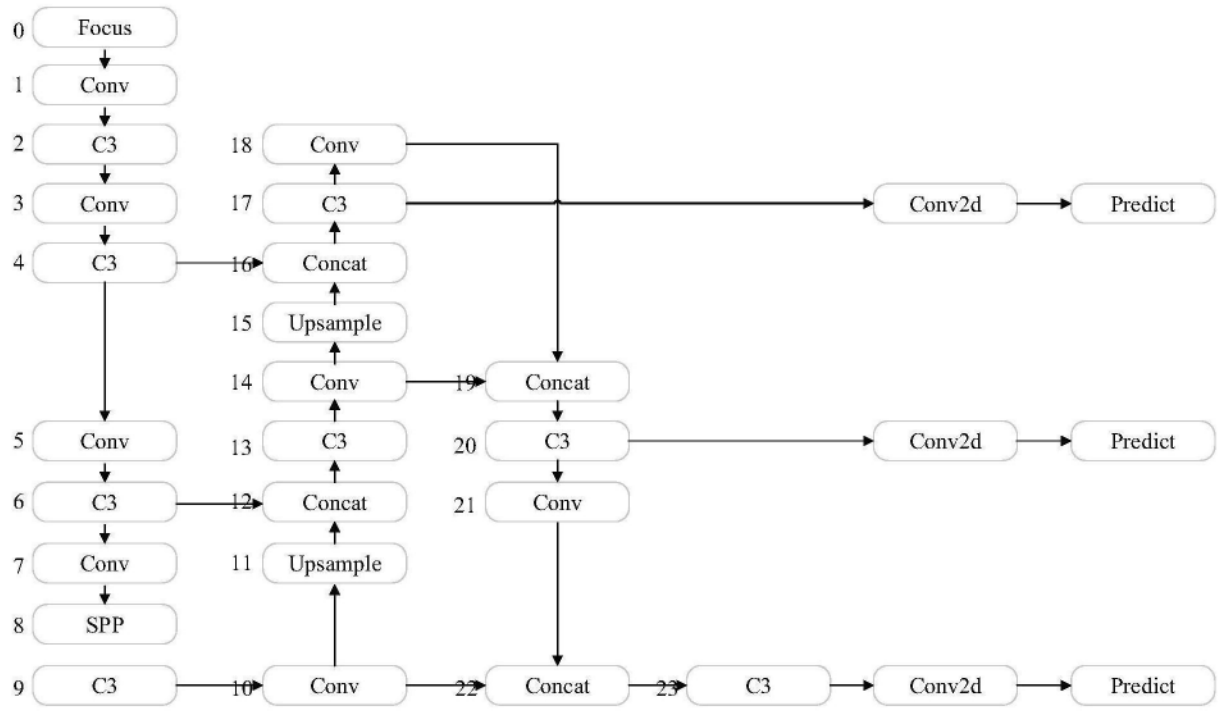


图3

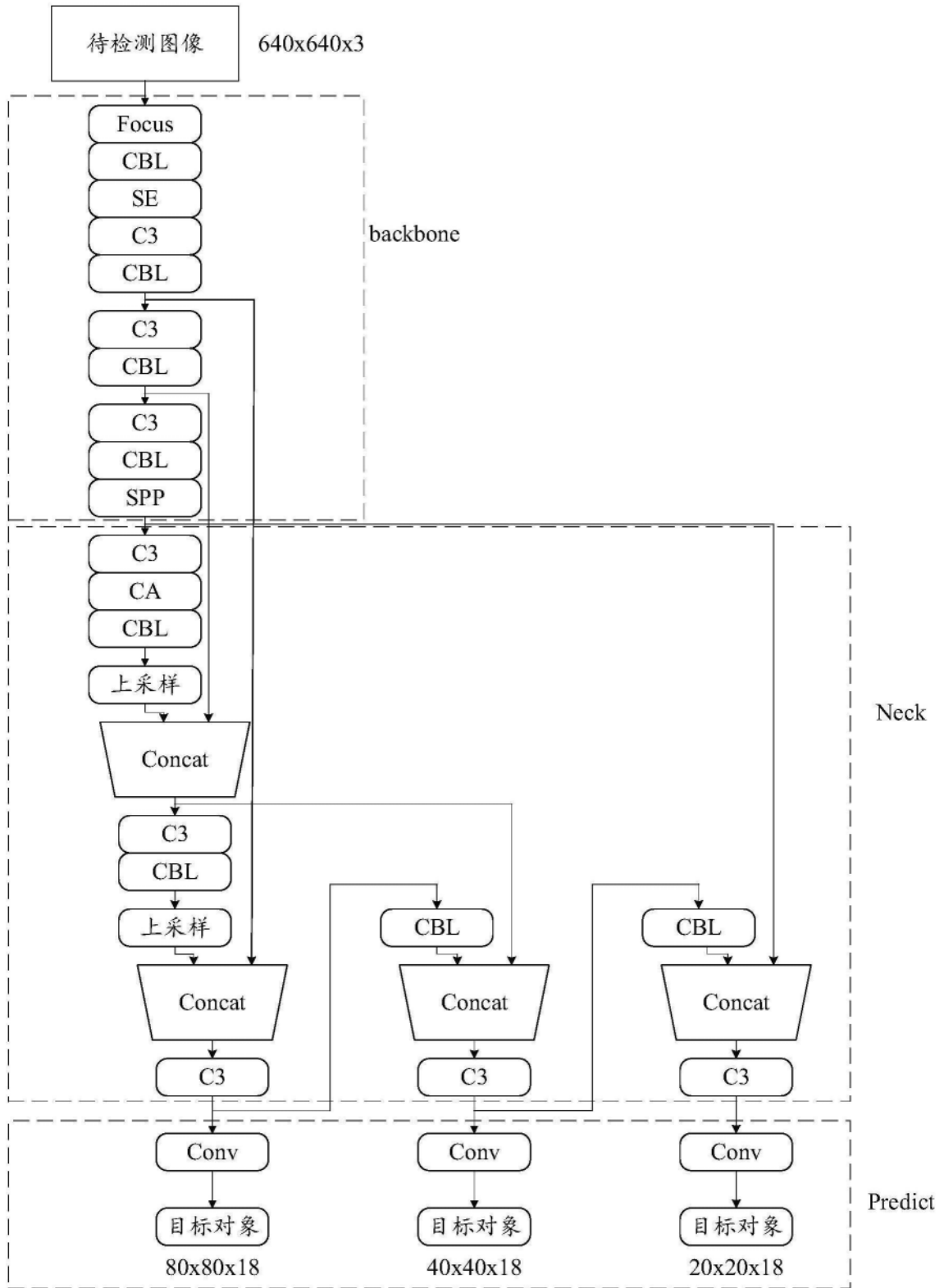


图4



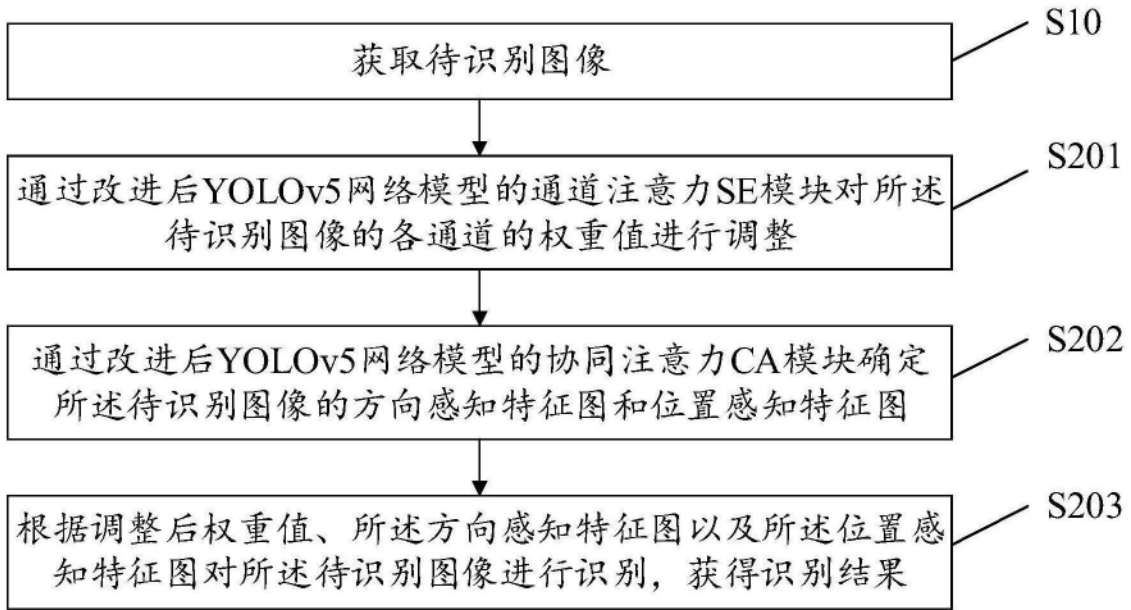


图5

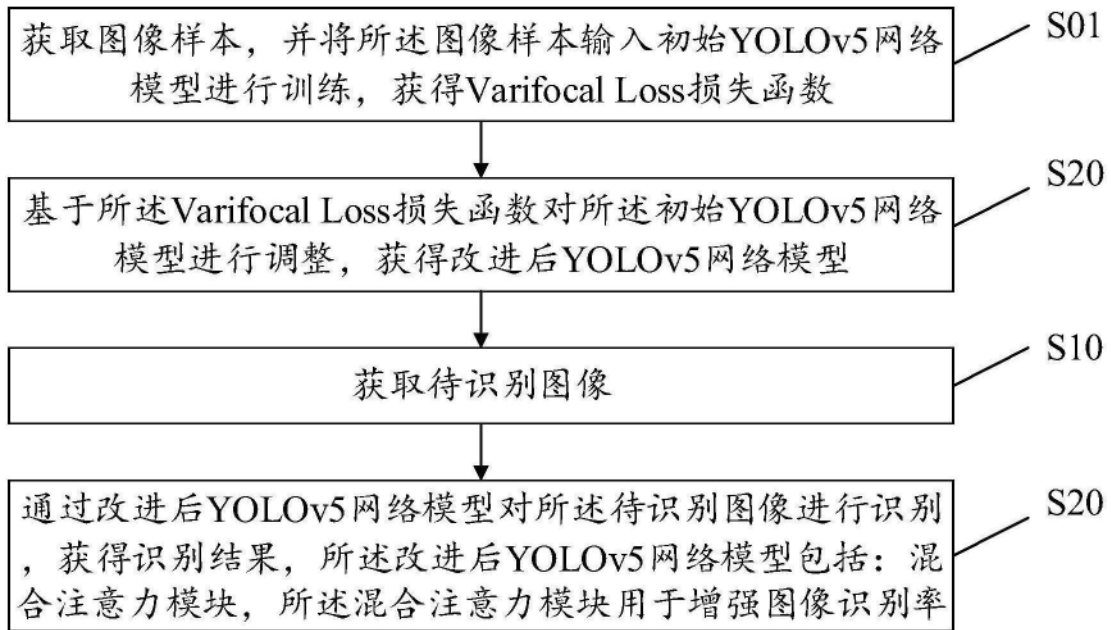


图6

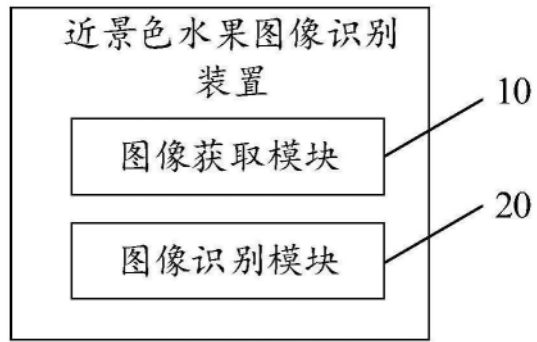


图7