



(12) 发明专利申请

(10) 申请公布号 CN 112085010 A

(43) 申请公布日 2020.12.15

(21) 申请号 202011173258.3

G06N 3/08 (2006.01)

(22) 申请日 2020.10.28

(71) 申请人 成都信息工程大学

地址 610015 四川省成都市西南航空港经
济开发区学府路1段24号

(72) 发明人 张朝龙 许源平 李婵玉 许志杰
黄健 卢军 谢攀 谭海龙
唐健峰

(74) 专利代理机构 北京正华智诚专利代理事务
所(普通合伙) 11870

代理人 李林合

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

权利要求书3页 说明书10页 附图3页

(54) 发明名称

一种基于图像识别的口罩检测和部署系统
及方法

(57) 摘要

本发明提供了一种基于图像识别的口罩检测
和部署系统及方法,属于生物识别技术领域,
包括:数据处理模块,用于分别收集人脸图像和
口罩图像,并分别对人脸图像和口罩图像进行标
注;模型训练模块,用于利用标注后的人脸图像
和口罩图像训练人脸口罩检测模型;模型推理模
块,用于根据人脸口罩检测模型,利用非极大值
抑制算法对人脸和口罩的位置进行检测;模型部
署模块,用于根据检测结果对人脸口罩检测模型
进行部署。本发明包括完整统一的图像数据采
集、数据标注、模型训练、模型部署与应用集成
的流程,其核心的多结构和多目标检测模型以及
自动化目标平台推理引擎部署特性可满足不同场
景、硬件配置、检测精度和人流量的需求。



1. 一种基于图像识别的口罩检测和部署系统,其特征在于,包括依次连接的数据处理模块、模型训练模块、模型推理模块以及模型部署模块;

所述数据处理模块,用于分别收集人脸图像和口罩图像,并分别对所述人脸图像和口罩图像进行标注;

所述模型训练模块,用于利用标注后的人脸图像和口罩图像训练人脸口罩检测模型;

所述模型推理模块,用于根据所述人脸口罩检测模型,利用非极大值抑制算法对人脸和口罩的位置进行检测;

所述模型部署模块,用于根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩检测和部署。

2. 一种基于图像识别的口罩检测和部署方法,其特征在于,包括以下步骤:

S1、数据处理:分别收集人脸图像和口罩图像,并分别对所述人脸图像和口罩图像进行标注;

S2、模型训练:利用标注后的人脸图像和口罩图像训练人脸口罩检测模型;

S3、模型推理:根据所述人脸口罩检测模型,利用非极大值抑制算法对人脸和口罩的位置进行检测;

S4、模型部署:根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩检测和部署。

3. 根据权利要求2所述的基于图像识别的口罩检测和部署方法,其特征在于,所述步骤S1包括以下步骤:

S101、分别收集人脸图像和口罩图像;

S102、合并所述人脸图像和口罩图像,并将合并后的图像按同比例进行分割,得到人脸类别图像和人脸佩戴口罩类别图像;

S103、将同比例分割后的剩余人脸图像划分为两部分,其中一部分作为人脸数据合并至所述人脸类别图像中,另外一部分利用在人脸脸上覆盖口罩的图片处理方法生成人脸佩戴口罩图像,合并至人脸佩戴口罩类别图像中;

S104、根据人脸类别图像和人脸佩戴口罩类别图像,利用标注工具分别标注人脸和口罩,并导出PASCAL VOC格式的XML标注文件,完成对人脸图像和口罩图像的标注。

4. 根据权利要求3所述的基于图像识别的口罩检测和部署方法,其特征在于,所述步骤S101中人脸图像的收集途径包括开源数据集、网络收集和非公开的监控数据;

所述口罩图像的收集途径包括:以现有公开的口罩检测数据集为基础,合并若干个口罩检测数据集;利用网络爬虫从网络上下载口罩图片;通过在人脸图像上制作佩戴口罩的口罩图片。

5. 根据权利要求4所述的基于图像识别的口罩检测和部署方法,其特征在于,所述步骤S2具体为:根据标注后的人脸图像和口罩图像,分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络构建第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型;其中,

所述YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络输入图片尺寸均设置为448×448,其输出类别均设置为2;

所述YOLOv3-tiny网络的训练达到的最大批次设置为6000,所述YOLO v3网络和YOLO

v4网络的训练达到的最大批次均设置为12000。

6. 根据权利要求5所述的基于图像识别的口罩检测和部署方法,其特征在于,所述根据标注后的人脸图像和口罩图像,分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络训练第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型,其包括以下步骤:

A1、利用YOLO预训练权重分别初始化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络;

A2、将标注后的人脸图像和口罩图像的尺寸缩放为448⁴448像素;

A 3、将缩放后的人脸图像和口罩图像分别分割成S²S个网格单元,其中,S表示整型常数;

A 4、在每个网格单元中分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取视觉特征,并分别设置YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的ReLU激活函数;

A 5、根据每个网格单元,分别利用YOLO v3网络的全连接层、YOLOv3-tiny网络的全连接层和YOLO v4网络的全连接层输出边界框 (x,y,w,h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$,其中, (x,y) 表示边界框的中心左边, w 和 h 分别表示边界框的宽度和长度, $class_i$ 表示第 i 个目标;

A 6、分别合并YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取所有网格单元的视觉特征以及对输出输出的边界框 (x,y,w,h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$,得到第一预测张量、第二预测张量以及第三预测张量;

A 7、根据所述第一预测张量、第二预测张量以及第三预测张量,分别计算得到第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度;

A 8、根据第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度,分别利用随机梯度下降法更新并优化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的参数,得到第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型,完成对人脸口罩检测模型的训练。

7. 根据权利要求6所述的基于图像识别的口罩检测和部署方法,其特征在于,所述步骤A 8中第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型的损失函数的表达式均如下:

$$\begin{aligned} & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(x_n - \hat{x}_j)^2 + (y_n - \hat{y}_j)^2 \right] \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[\left(\sqrt{w_n} - \sqrt{\hat{w}_j} \right)^2 + \left(\sqrt{h_n} - \sqrt{\hat{h}_j} \right)^2 \right] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} I_i^{\text{obj}} \sum_{m \in \text{classes}} (p_n(m) - \hat{p}_j(m))^2 \end{aligned}$$

其中, λ_{coord} 表示对预测的中心坐标和边界框的宽度做损失的权重, S^2 表示网格单元

的总数, B 表示预测边界框的总数, I_{ij}^{obj} 表示当第 i 个网格单元存在目标则第 j 个边界框预测值对该预测值有效, i 表示网格单元的个数, j 表示边界框的个数, x_n 表示第 n 个预测边界框的 x 轴坐标值, \hat{x}_j 表示训练数据中第 j 个边界框实际的 x 轴坐标值, y_n 表示第 n 个预测边界框的 y 轴坐标值, \hat{y}_j 表示训练数据中第 j 个边界框实际的 y 轴坐标值, w_n 表示第 n 个预测边界框的宽度, \hat{w}_j 表示训练数据中第 j 个边界框实际的宽度, h_n 表示第 n 个预测边界框的高度, \hat{h}_j 表示训练数据中第 j 个边界框实际的高度, C_i 表示目标的置信度, \hat{C}_i 表示预测的边界框与实际的边界框的交叉部分比率, λ_{noobj} 表示对置信度预测的损失权重, I_{ij}^{noobj} 表示第 i 个网格单元不存在目标则第 j 个边界框预测值对该预测值有效, I_i^{obj} 表示第 i 个网格单元有目标, $p_n(m)$ 表示第 n 个预测边界框的目标类别, $\hat{p}_j(m)$ 表示第 j 个边界框的实际目标类别, c 表示目标, $classes$ 表示目标集合。

8. 根据权利要求7所述的基于图像识别的口罩检测和部署方法, 其特征在于, 所述步骤 S3 包括以下步骤:

S301、根据所述第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型, 分别利用非极大值抑制算法对目标候选框进行去重处理, 得到第一目标区域、第二目标区域以及第三目标区域;

S302、根据所述第一目标区域、第二目标区域以及第三目标区域, 分别利用预设的概率阈值检测图像中人脸和口罩出现的位置, 完成对人脸和口罩位置的检测。

9. 根据权利要求8所述的基于图像识别的口罩检测和部署方法, 其特征在于, 所述步骤 S4 包括以下步骤:

S401、分别搭建第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型的转换环境, 并分别安装编程语言运行库 Python、开放式神经网络交换运行库 onnx 以及开发套件 RKNN Toolkit;

S402、利用编程语言运行库 Python 和开放式神经网络交换运行库 onnx 分别读取第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型的 cfg 和 $weights$ 文件, 分别导出第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型中的 ONNX 结构文件;

S403、利用 RKNN Toolkit 分别将第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型中的 ONNX 结构文件转换为 RKNN 模型文件;

S404、利用 RKNN API 编写应用程序加装 RKNN 模型文件, 分别进行第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型的推理应用;

S405、根据推理应用结果编译应用程序, 并将应用程序和第一人臉口罩检测模型、第二人脸口罩检测模型以及第三人臉口罩检测模型的 RKNN 模型文件分别部署至硬件平台;

S406、读取摄像头的监控图像, 并利用 RKNN 模型文件调用推理引擎对监控图像进行分析, 并输出分析结果, 完成基于图像识别的口罩检测和部署。

一种基于图像识别的口罩检测和部署系统及方法

技术领域

[0001] 本发明属于生物识别技术领域,尤其涉及一种基于图像识别的口罩检测和部署系统及方法。

背景技术

[0002] 口罩作为一种常用的医用和卫生用品,可有效阻挡有害气体、飞沫和病毒进入佩戴者的口鼻,从而可以降低口罩佩戴者的疾病感染风险。正确佩戴口罩在对新冠病毒疫情和其它传染病的防控起到了重要的作用,有效地减少了病毒感染者并降低了传染病毒的扩散速度与范围。因此,在公共场所部署口罩佩戴检测装置和系统,识别行人是否正确佩戴了口罩,并进行提醒和预警,在疫情防控中起到了积极的作用。比如,在地铁口等人流密集区域部署口罩检测佩戴检测系统,当检测到有乘客未佩戴口罩就进入地铁站时,及时通知安保人员,并提醒乘客佩戴口罩,从而可以确保行人进入密集区域之前佩戴口罩,以实现有效的疫情防控措施。

[0003] 目前主要的人脸佩戴口罩检测技术是人工检测,即有工作人员站在检测站(如地铁口、小区出入口等),通过人眼观察行人面部来确定行人是否佩戴了口罩,并对未戴口罩的行人进行口头提醒和警告等。人工检测方法造成了相关工作人员的劳动强度大,并容易出现漏检现象,同时也把工作人员暴露在易感染的危险环境中。此外,也有通过在口罩中嵌入传感器装置,该传感器在接触到皮肤后可呈现不同的状态变化,比如可以通过静脉检测来判断是否与皮肤接触。该装置通过实时读取传感器数据,来判断是否被佩戴,外部设备读取口罩上的传感器数据,来确定行人是否佩戴了口罩。入侵式检测方法需要改变口罩的原有结构,导致成本增加,影响佩戴的舒适性,并且也难以应用推广。

[0004] 随着计算机视觉技术的发展,人脸检测和识别技术得到了飞速发展和广泛应用,因此也随之出现了基于图像识别技术的口罩检测。这种方法首先需要定位在图像上定位人脸位置,然后分析人脸属性。人脸属性分析又包括两种:1)检测人脸器官,如检测眼睛、鼻子和嘴的位置,当检测到人脸却检测不到嘴时,可以认为是佩戴了口罩;2)颜色统计,由于人脸皮肤和口罩的颜色不同,通过在检测到人脸后,统计人脸区域的颜色特征,如生成颜色直方图(Color Histogram)和方向梯度直方图(Histogram of Oriented Gradients,HOG),然后使用机器学习算法(如SVM(Support Vector Machine))进行分类,识别该人脸是否为佩戴口罩的人脸。基于人脸检测和人脸属性分析的方法识别率不高,其缺点主要表现在:1)当人脸被口罩覆盖后,本身就会导致人脸检测算法的效果不佳。由于检测不到人脸,则无法进行后续的口罩检测;2)以是否在人脸上检测到嘴的结论来判断是否佩戴口罩的方法容易出现误判的情况,如侧脸、用手或其它物体覆盖面部等情况都会被识别为佩戴了口罩;3)由于人的皮肤颜色差异以及使用具有与口罩类似颜色的物体覆盖面部时,统计人脸区域颜色特征的方法也同样会出现误判的情况,这些问题都会导致误检和漏检的情况。

发明内容

[0005] 针对现有技术中的上述不足,本发明提供一种基于图像识别的口罩检测和部署系统及方法,解决了上述的问题。

[0006] 为了达到以上目的,本发明采用的技术方案为:

本方案提供了一种基于图像识别的口罩检测和部署系统,包括依次连接的数据处理模块、模型训练模块、模型推理模块以及模型部署模块;

所述数据处理模块,用于分别收集人脸图像和口罩图像,并分别对所述人脸图像和口罩图像进行标注;

所述模型训练模块,用于利用标注后的人脸图像和口罩图像训练人脸口罩检测模型;

所述模型推理模块,用于根据所述人脸口罩检测模型,利用非极大值抑制算法对人脸和口罩的位置进行检测;

所述模型部署模块,用于根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩检测和部署。

[0007] 基于上述系统,本发明还提供了一种基于图像识别的口罩检测和部署方法,包括以下步骤:

S1、数据处理:分别收集人脸图像和口罩图像,并分别对所述人脸图像和口罩图像进行标注;

S2、模型训练:利用标注后的人脸图像和口罩图像训练人脸口罩检测模型;

S3、模型推理:根据所述人脸口罩检测模型,利用非极大值抑制算法对人脸和口罩的位置进行检测;

S4、模型部署:根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩检测和部署。

[0008] 进一步地,所述步骤S1包括以下步骤:

S101、分别收集人脸图像和口罩图像;

S102、合并所述人脸图像和口罩图像,并将合并后的图像按同比例进行分割,得到人脸类别图像和人脸佩戴口罩类别图像;

S103、将同比例分割后的剩余人脸图像划分为两部分,其中一部分作为人脸数据合并至所述人脸类别图像中,另外一部分利用在人脸脸上覆盖口罩的图片处理方法生成人脸佩戴口罩图像,合并至人脸佩戴口罩类别图像中;

S104、根据人脸类别图像和人脸佩戴口罩类别图像,利用标注工具分别标注人脸和口罩,并导出PASCAL VOC格式的XML标注文件,完成对人脸图像和口罩图像的标注。

[0009] 再进一步地,所述步骤S101中人脸图像的收集途径包括开源数据集、网络收集和非公开的监控数据;

所述口罩图像的收集途径包括:以现有公开的口罩检测数据集为基础,合并若干个口罩检测数据集;利用网络爬虫从网络上下载口罩图片;通过在人脸图像上制作佩戴口罩的口罩图片。

[0010] 再进一步地,所述步骤S2具体为:根据标注后的人脸图像和口罩图像,分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络构建第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型;其中,

所述YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络输入图片尺寸均设置为 448×448 ，其输出类别均设置为2；

所述YOLOv3-tiny网络的训练达到的最大批次设置为6000，所述YOLO v3网络和YOLO v4网络的训练达到的最大批次均设置为12000。

[0011] 再进一步地，所述根据标注后的人脸图像和口罩图像，分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络训练第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型，其包括以下步骤：

A1、利用YOLO预训练权重分别初始化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络；

A2、将标注后的人脸图像和口罩图像的尺寸缩放为 448×448 ；

A 3、将缩放后的人脸图像和口罩图像分别分割成 $S \times S$ 个网格单元，其中， S 表示整型常数；

A 4、在每个网格单元中分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取视觉特征，并分别设置YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的ReLU激活函数；

A 5、根据每个网格单元，分别利用YOLO v3网络的全连接层、YOLOv3-tiny网络的全连接层和YOLO v4网络的全连接层输出边界框 (x, y, w, h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$ ，其中， (x, y) 表示边界框的中心左边， w 和 h 分别表示边界框的宽度和长度， $class_i$ 表示第 i 个目标；

A 6、分别合并YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取所有网格单元的视觉特征以及对应输出的边界框 (x, y, w, h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$ ，得到第一预测张量、第二预测张量以及第三预测张量；

A 7、根据所述第一预测张量、第二预测张量以及第三预测张量，分别计算得到第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度；

A 8、根据第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度，分别利用随机梯度下降法更新并优化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的参数，得到第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型，完成对人脸口罩检测模型的训练。

[0012] 再进一步地，所述步骤A 8中第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型的损失函数的表达式均如下：

$$\begin{aligned} & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(x_n - \hat{x}_j)^2 + (y_n - \hat{y}_j)^2 \right] \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[\left(\sqrt{w_n} - \sqrt{\hat{w}_j} \right)^2 + \left(\sqrt{h_n} - \sqrt{\hat{h}_j} \right)^2 \right] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (c_i - \hat{c}_i)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} (c_i - \hat{c}_i)^2 + \sum_{i=0}^{S^2} I_i^{\text{obj}} \sum_{m \in \text{classes}} (p_n(m) - \hat{p}_j(m))^2 \end{aligned}$$

其中， λ_{coord} 表示对预测的中心坐标和边界框的宽度做损失的权重， S^2 表示网格单元的总数， B 表示预测边界框的总数， I_{ij}^{obj} 表示当第 i 个网格单元存在目标则第 j 个边界框预测值对

该预测值有效, i 表示网格单元的个数, j 表示边界框的个数, x_n 表示第 n 个预测边界框的 x 轴坐标值, \hat{x}_j 表示训练数据中第 j 个边界框实际的 x 轴坐标值, y_n 表示第 n 个预测边界框的 y 轴坐标值, \hat{y}_j 表示训练数据中第 j 个边界框实际的 y 轴坐标值, w_n 表示第 n 个预测边界框的宽度, \hat{w}_j 表示训练数据中第 j 个边界框实际的宽度, h_n 表示第 n 个预测边界框的高度, \hat{h}_j 表示训练数据中第 j 个边界框实际的高度, C_i 表示目标的置信度, \hat{C}_i 表示预测的边界框与实际的边界框的交叉部分比率, λ_{noobj} 表示对置信度预测的损失权重, I_{ij}^{noobj} 表示第 i 个网格单元不存在目标则第 j 个边界框预测值对该预测值有效, I_i^{obj} 表示第 i 个网格单元有目标, $p_n(m)$ 表示第 n 个预测边界框的目标类别, $\hat{p}_j(m)$ 表示第 j 个边界框的实际目标类别, m 表示目标, $classes$ 表示目标集合。

[0013] 再进一步地, 所述步骤S3包括以下步骤:

S301、根据所述第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型, 分别利用非极大值抑制算法对目标候选框进行去重处理, 得到第一目标区域、第二目标区域以及第三目标区域;

S302、根据所述第一目标区域、第二目标区域以及第三目标区域, 分别利用预设的概率阈值检测图像中人脸和口罩出现的位置, 完成对人脸和口罩位置的检测。

[0014] 再进一步地, 所述步骤S4包括以下步骤:

S401、分别搭建第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型的转换环境, 并分别安装编程语言运行库Python、开放式神经网络交换运行库onnx以及开发套件RKNN Toolkit;

S402、利用编程语言运行库Python和开放式神经网络交换运行库onnx分别读取第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型的cfg和weights文件, 分别导出第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型中的ONNX结构文件;

S403、利用RKNN Toolkit分别将第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型中的ONNX结构文件转换为RKNN模型文件;

S404、利用RKNN API编写应用程序加装RKNN模型文件, 分别进行第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型的推理应用;

S405、根据推理应用结果编译应用程序, 并将应用程序和第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型的RKNN模型文件分别部署至硬件平台;

S406、读取摄像头的监控图像, 并利用RKNN模型文件调用推理引擎对监控图像进行分析, 并输出分析结果, 完成基于图像识别的口罩检测和部署。

[0015] 本发明的有益效果:

(1) 本发明提供了一个基于深度学习和目标检测的人脸口罩检测与部署方法, 实现从监控视频和图像中自动识别行人、检测人脸, 并检测人脸是否佩戴了口罩。本发明包括完整的统一的图像数据采集、数据标注、模型训练、模型部署与应用集成的流程, 可实现口罩检测的全栈开发流程, 方便快速地把人脸口罩检测模型部署到目标平台, 其核心的多结构和多目标检测模型, 以及自动化目标平台推理引擎部署特性可满足不同场景、硬件配置、检测精度和人流量的需求。

[0016] (2) 本发明使用软硬件系统替代人工操作：本发明可用于开发智能的口罩检测软硬件系统，以替代人工观察的检测方式，本发明能有效地减少工作人员的劳动强度，提高检测效率和精度，并可保护相关工作人员的健康和安全。

[0017] (3) 本发明解决了口罩检测数据集不足的问题：生物识别应用中主要以人脸数据集为主，基于这些数据集研发了多种基于人脸检测和人脸识别的应用，但是口罩检测和识别的数据集非常少，很难支持构建高精度的口罩检测算法模型。本发明采用多种方式增加口罩数据集，为训练基于深度学习的口罩检测模型提供了数据支持。

[0018] (4) 本发明提高了口罩检测的正确率：本发明以YOLO为基础构建人脸和口罩检测模型，YOLO在目标检测应用领域中具有非常好的性能和效果。本发明采用迁移学习的方法在预训练的YOLO模型之上构建人脸和口罩检测模型，通过引入loss函数和随机梯度算法，确保深度卷积神经网络在训练数据集中收敛，同时防止其出现过拟合现象，从而得到高精度的人脸和口罩检测模型。此外，通过引入NMS，生成最终的识别结果。

[0019] (5) 本发明降低了部署成本和难度：传统部署深度学习模型的方法通常都需要使用GPU计算硬件，对计算资源要求较高，为满足在低功耗的嵌入式设备上部署该人脸和口罩检测模型，本发明设计了一种模型部署方法，通过把一个较大的模型转换为ONNX中间结构，然后使用目标平台工具进行转换和优化，从而使该人脸和口罩检测模型可以在嵌入式和边缘计算设备上部署运行，降低了对计算资源的需求和成本。

附图说明

[0020] 图1为本发明的系统结构示意图。

[0021] 图2为本发明的方法流程图。

[0022] 图3为本实施例中人脸口罩检测模型的部署流程示意图。

具体实施方式

[0023] 下面对本发明的具体实施方式进行描述，以便于本技术领域的技术人员理解本发明，但应该清楚，本发明不限于具体实施方式的范围，对本技术领域的普通技术人员来讲，只要各种变化在所附的权利要求限定和确定的本发明的精神和范围内，这些变化是显而易见的，一切利用本发明构思的发明创造均在保护之列。

[0024] 实施例1

目前，公共场所都安装有监控摄像头。在某些区域（如办公写字楼和学校等）入口处甚至安装了人脸识别设备。采用图像识别和深度学习技术，通过从这些监控图像中自动分析行人是否佩戴口罩，并自动触发警报提醒，可有效解决目前口罩检测的难题。此外，通过构建统一的图像数据采集、数据标注、模型训练、模型部署与应用集成的流程，可实现口罩检测的全栈开发流程，方便快速地把人脸口罩检测模型部署到目标平台。

[0025] 如图1所示，本发明提供了一种基于图像识别的口罩检测和部署系统，包括依次连接的数据处理模块、模型训练模块、模型推理模块以及模型部署模块；数据处理模块，用于分别收集人脸图像和口罩图像，并分别对所述人脸图像和口罩图像进行标注；模型训练模块，用于利用标注后的人脸图像和口罩图像训练人脸口罩检测模型；模型推理模块，用于根据所述人脸口罩检测模型，利用非极大值抑制算法对人脸和口罩的位置进行检测；模型部

署模块,用于根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩检测和部署。

[0026] 本实施例中,在进行部署后,部署监控结果可进一步处理:如集成声音提醒等。

[0027] 本实施例中,本发明人脸口罩检测部署方法包括图像数据处理、模型训练、模型推理、模型部署四个部分,首先是收集图像数据,主要包括人脸和口罩的图像数据,同时还需要对图像中出现人脸和口罩的位置进行标注;在标注完数据后,就需要使用这些数据训练目标检测模型,从而得到人脸口罩检测模型;该模型只是所使用训练框架的特定模型,不具有通用性,为了实现在目标平台和系统上运行,还需要进行模型的移植和部署。

[0028] 本实施例中,本发明实现从监控视频和图像中自动识别行人、检测人脸,并检测人脸是否佩戴了口罩。本发明包括完整的统一的图像数据采集、数据标注、模型训练、模型部署与应用集成的流程,可实现口罩检测的全栈开发流程,方便快速地把人脸口罩检测模型部署到目标平台,其核心的多结构和多目标检测模型,以及自动化目标平台推理引擎部署特性可满足不同场景、硬件配置、检测精度和人流量的需求。

[0029] 实施例2

如图2所示,本发明提供了一种基于图像识别的口罩检测和部署方法,其实现方法如下:

S1、数据处理:分别收集人脸图像和口罩图像,并分别对所述人脸图像和口罩图像进行标注,其实现方法如下:

S101、分别收集人脸图像和口罩图像;

S102、合并所述人脸图像和口罩图像,并将合并后的图像按同比例进行分割,得到人脸类别图像和人脸佩戴口罩类别图像;

S103、将同比例分割后的剩余人脸图像划分为两部分,其中一部分作为人脸数据合并至所述人脸类别图像中,另外一部分利用在人脸上覆盖口罩的图片处理方法生成人脸佩戴口罩图像,合并至人脸佩戴口罩类别图像中;

S104、根据人脸类别图像和人脸佩戴口罩类别图像,利用标注工具分别标注人脸和口罩,并导出PASCAL VOC格式的XML标注文件,完成对人脸图像和口罩图像的标注。

[0030] 本实施例中,本发明所使用的图像数据包括两种用途:1)用于训练模型的训练和测试图像数据;2)用于在应用中检测人脸和口罩的图像数据。

[0031] 本实施例中,在应用中本发明从监控图像中检测并识别人脸和口罩目标。图像的来源为摄像头,本发明通过UVC协议从USB摄像头中读取视频流,或者使用RTSP协议从IP摄像头中读取视频流,使用ffmpeg视频分析库从视频流中提取视频关键帧,然后使用在目标平台部署的推理模型进行图像分析,检测画面中的行人和人脸,以及行人是否佩戴口罩。

[0032] 本实施例中,用于训练模型的图片数据采用离线收集和人工标注的方式,图片来源主要为开源数据集、网络收集和企业非公开的监控数据。目前开源的人脸检测和识别的数据集比较多,如FDDB(Face Detection Data Set and Benchmark)、Wider-face和MALF等人脸检测数据集。本发明主要使用这些数据集作为人脸检测的图像数据。而口罩检测的数据集目前较少,现有开源的口罩检测数据集规模也较小,并且标注也不规范,为获取足够多的口罩检测训练数据,本方法采用如下策略收集口罩图片数据:以现有公开的口罩检测数据集为基础,合并多个口罩检测数据集;使用网络爬虫从网络上下载图片,如使用关键字

(如“口罩”、“佩戴口罩”、“防疫”、“疫情”、“facemask”等)从百度图片、淘宝、微博等网络平台搜索并下载相关图片,然后人工筛选符合条件的图片;通过在人脸数据集中的人脸图片上覆盖口罩图片,来制作佩戴口罩的图片。在这种方法中,首先使用人脸检测算法或者从标注数据中(若有)检测出人脸的位置和大小,然后对口罩图片进行比例变化,并把口罩图片添加到人脸下方,确保口罩覆盖人脸下方1/2至2/3的区域。

[0033] 本实施例中,当完成人脸和口罩数据集采集后,本发明需要合并上述两种图像数据,并对图像数据进行标注。合并人脸图像和口罩图像,并分为两个类别:人脸和口罩(人脸佩戴口罩),两个类别的图片数据均等分割,由于口罩数据集的图片较少,此时将剩余大量的人脸图片。对剩余的人脸图片再分为两等份,其中一份作为人脸数据合并至人脸类别中;另一部分图片则采用在人脸上覆盖口罩图片的方法人工生成佩戴口罩的图片数据,以增强口罩检测数据集,使用Colabeler标注工具标注图片上出现的人脸和口罩,标注框为矩形,该标注框即为ground truth,标注完成后,导出PASCAL VOC格式的XML标注文件。

[0034] S2、模型训练:利用标注后的人脸图像和口罩图像训练人脸口罩检测模型,其具体为:

根据标注后的人脸图像和口罩图像,分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络构建第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型;YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络输入图片尺寸均设置为 448×448 ,其输出类别均设置为2;YOLOv3-tiny网络的训练达到的最大批次设置为6000,所述YOLO v3网络和YOLO v4网络的训练达到的最大批次均设置为12000,其实现方法如下:

A1、利用YOLO预训练权重分别初始化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络;

A2、将标注后的人脸图像和口罩图像的尺寸缩放为 448×448 ;

A 3、将缩放后的人脸图像和口罩图像分别分割成 $S \times S$ 个网格单元,其中, S 表示整型常数;

A 4、在每个网格单元中分别利用YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取视觉特征,并分别设置YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的ReLU激活函数;

A 5、根据每个网格单元,分别利用YOLO v3网络的全连接层、YOLOv3-tiny网络的全连接层和YOLO v4网络的全连接层输出边界框 (x,y,w,h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$,其中, (x,y) 表示边界框的中心左边, w 和 h 分别表示边界框的宽度和长度, $class_i$ 表示第 i 个目标;

A 6、分别合并YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络提取所有网格单元的视觉特征以及对应输出的边界框 (x,y,w,h) 、包含目标的置信度 $Pr(object)$ 以及每个类别的预测条件概率 $Pr(class_i|object)$,得到第一预测张量、第二预测张量以及第三预测张量;

A 7、根据所述第一预测张量、第二预测张量以及第三预测张量,分别计算得到第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度;

A 8、根据第一卷积神经网络的梯度、第二卷积神经网络的梯度以及第三卷积神经网络的梯度,分别利用随机梯度下降法更新并优化YOLO v3网络、YOLOv3-tiny网络和YOLO v4网络的参数,得到第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩检测模型,完成对人脸口罩检测模型的训练。

[0035] 本实施例中,第一人脸口罩检测模型、第二人脸口罩检测模型以及第三人脸口罩

$$\begin{aligned}
& \text{检测模型的损失函数的表达式均如下: } \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [(x_n - \hat{x}_j)^2 + (y_n - \hat{y}_j)^2] \\
& + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(\sqrt{w_n} - \sqrt{\hat{w}_j})^2 + (\sqrt{h_n} - \sqrt{\hat{h}_j})^2 \right] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (c_i - \hat{c}_i)^2 \\
& + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} (c_i - \hat{c}_i)^2 + \sum_{i=0}^{S^2} I_i^{\text{obj}} \sum_{m \in \text{classes}} (p_n(m) - \hat{p}_j(m))^2
\end{aligned}$$

其中, λ_{coord} 表示对预测的中心坐标和边界框的宽度做损失的权重, S^2 表示网格单元的总数, B 表示预测边界框的总数, I_{ij}^{obj} 表示当第 i 个网格单元存在目标则第 j 个边界框预测值对该预测值有效, i 表示网格单元的个数, j 表示边界框的个数, x_n 表示第 n 个预测边界框的 x 轴坐标值, \hat{x}_j 表示训练数据中第 j 个边界框实际的 x 轴坐标值, y_n 表示第 n 个预测边界框的 y 轴坐标值, \hat{y}_j 表示训练数据中第 j 个边界框实际的 y 轴坐标值, w_n 表示第 n 个预测边界框的宽度, \hat{w}_j 表示训练数据中第 j 个边界框实际的宽度, h_n 表示第 n 个预测边界框的高度, \hat{h}_j 表示训练数据中第 j 个边界框实际的高度, c_i 表示目标的置信度, \hat{c}_i 表示预测的边界框与实际的边界框的交叉部分比率, λ_{noobj} 表示对置信度预测的损失权重, I_{ij}^{noobj} 表示第 i 个网格单元不存在目标则第 j 个边界框预测值对该预测值有效, I_i^{obj} 表示第 i 个网格单元有目标, $p_n(m)$ 表示第 n 个预测边界框的目标类别, $\hat{p}_j(m)$ 表示第 j 个边界框的实际目标类别, m 表示目标, classes 表示目标集合。

[0036] 本实施例中, 本发明以Darknet深度学习库和YOLO神经网络架构构建人脸口罩检测模型, YOLO是一种基于深度学习的目标检测模型, 可以检测场景的物体, 如人、汽车和狗等物体目标, 但是YOLO不能直接用于口罩检测。因此, 本发明基于YOLO构建人脸和口罩的检测模型, 本发明搭建了三种不同的YOLO网络架构, 包括YOLO v3、YOLOv3-tiny和YOLO v4网络, 然后设置输入图片尺寸为 448×448 , 输出类别为2, YOLOv3-tiny的max batches为6000, YOLO v3和YOLO v4的max batches为12000, 启用GPU, 然后分别训练这三个网络模型: 首先使用预训练的权重初始化YOLO网络, 预训练的权重可以确保视觉特征提取的有效性和正确性, 从而可以满足迁移学习的条件; 对于一张待训练的图片, 其包含图片本身和标准目标类别(人脸和口罩)的区域(ground truth), 把图片的尺寸缩放为 448×448 像素; 把缩放后的图片分割为 $S \times S$ 个网格单元; 每一个单元网络使用多层CNN深度神经网络提取视觉特征, 并且使用Leaky ReLU激活函数, 以避免使用ReLU函数时某些网格单元无法激活的问题; 每个单元网络最后使用线性全连接层输出一个边界框(bounding box), 其值为 (x, y, w, h) , 一个包含目标的置信度以及对每个类别的预测条件概率; 对每个单元网络完成特征提取和输出后, 合并所有单元网络的输出, 形成一个预测张量; 根据该图片的标注数据计算卷积神经网络的梯度, 其中边界框的准确度采用IOU来表示(Intersection Over Union), 即目标置信度为: $Pr(\text{object}) = IOU_{\text{pred}}^{\text{truth}}$, 其中, truth即为实际的边界框, pred为预测的边界框, 最后使用随机梯度下降法, 对模型的参数进行更新, 优化YOLO网络检测模型, 从而得到

人脸口罩检测模型。

[0037] 本实施例中,在训练过程中,通过观察MAP分析检测模型的学习效果,并保留检测效果最高的权重作为已训练的检测模型。经过上述的模型训练过程之后,则可获得已训练的人脸口罩检测模型文件(cfg和weights)。

[0038] S3、模型推理:根据所述人脸口罩检测模型,利用非极大值抑制算法对人脸和口罩的位置进行检测,其实现方法如下:

S301、根据所述第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型,分别利用非极大值抑制算法对目标候选框进行去重处理,得到第一目标区域、第二目标区域以及第三目标区域;

S302、根据所述第一目标区域、第二目标区域以及第三目标区域,分别利用预设的概率阈值检测图像中人脸和口罩出现的位置,完成对人脸和口罩位置的检测。

[0039] 本实施例中,经过上述训练方法在人脸和口罩数据集训练后,本发明可得到一个已训练的人脸口罩检测模型,该模型可在Darknet深度学习框架下运行。从上述的训练流程可知,该模型只能输出一个预测张量,还无法直接输出人脸和口罩的检测结果。上传人脸口罩检测模型输出的预测张量实际上表示一张图片上的多个候选框,以及每个候选框所属目标类别的置信度,对于同一个目标(人脸或口罩),可能会出现多个相交的候选框,即一个目标会被多次检测到,为定位每一个目标的边界和类别,本发表使用非极大值抑制算法(non maximum suppression, NMS)进行目标候选框去重,最后可得到置信度最大的非重叠检测框,即为目标区域,该框的置信概率即为目标的概率,通过设定一个概率阈值,则可检测图像中的人脸和口罩出现的位置,从而可以确定行人是否佩戴了口罩,即当检测到独立的人脸时,则说明该人脸所对应的行人未佩戴口罩。

[0040] S4、模型部署:根据检测结果对人脸口罩检测模型进行部署,完成基于图像识别的口罩部署,其实现方法如下:

S401、分别搭建第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型的转换环境,并分别安装编程语言运行库Python、开放式神经网络交换运行库onnx以及瑞芯微提供的开发套件RKNN Toolkit;

S402、利用编程语言运行库Python和开放式神经网络交换运行库onnx分别读取第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型的cfg和weights文件,分别导出第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型中的ONNX结构文件;

S403、利用RKNN Toolkit分别将第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型中的ONNX结构文件转换为RKNN模型文件;

S404、利用RKNN API编写应用程序加装RKNN模型文件,分别进行第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型的推理应用;

S405、根据推理应用结果编译应用程序,并将应用程序和第一人人脸口罩检测模型、第二人脸口罩检测模型以及第三人人脸口罩检测模型的RKNN模型文件分别部署至硬件平台;

S406、读取摄像头的监控图像,并利用RKNN模型文件调用推理引擎对监控图像进行分析,并输出分析结果,完成基于图像识别的口罩检测和部署。

[0041] 本实施例中,训练后的人脸口罩检测模型只能使用Darknet进行部署和使用。

Darknet是一个使用C语言实现的深度学习框架,包含训练模型和推理模型,并集成了常用的算子。使用Darknet部署本发明的人脸口罩检测模型在一定程度上可以满足应用需求,如在PC版(包括Windows和Linux操作系统)的应用程序中集成Darknet,加载本发明的人脸口罩检测模型,从而实现上位机的口罩佩戴检测功能。但是,由于Darknet对框架较为复杂,对计算资源要求也相对较高(如通常都需要在PC机上安装GPU,以支持卷积神经网络的运算),这种方法在计算能力相对较低的设备上部署时难以支持实时分析,尤其是嵌入式设备和边缘计算设备,而在部署应用时,训练引擎已经不再被需要,只需要使用推理引擎即可。因此,本发明设计一个人脸口罩检测模型的在嵌入式设备和边缘计算设备的部署方式,如图3所示,已训练模型包括网络配置文件(cfg)和参数权重文件(weights),首先使用Python和onnx库把cfg和weights文件转化为ONNX模型(中间模型)结构文件,ONNX是深度学习框架中迁移模型的中间表达格式框架,它不依赖于具体的深度学习框架,表达深度学习模型的网络结构和权重;在得到onnx结构文件后,再使用目标平台的生成工具(目标平台Toolkits),生成对应平台的网络模型(如图3中的目标模型1、目标模型2和目标模型3)和权重,以开发目标平台的应用(如图3中的应用1、应用2和应用3)。如可以使用RKNN-Toolkit则可以把onnx文件转换为RKNN模型,从而实现在瑞芯微的NPU上部署运行;也可以把onnx转化为caffe结构模型,然后使用海思的Ruyi Studio工具把caffe结构模型转化为海思的NNIE结构文件,以支持在海思的NPU上运行。此外,也可以直接在支持onnx的推理引擎上加载并运行onnx模型,如onnxruntime和OpenCV DNN。

[0042] 本实施例中,以瑞芯微的RK3399 Pro硬件平台说明如何部署:

首先在PC机上搭建模型转换环境,需要安装Python和onnx库,同时也需要安装RKNN Toolkit;使用Python的onnx库读取cfg和weights文件,然后导出为ONNX结构文件,如YOLOv3-tiny.onnx;使用RKNN Toolkit把ONNX结构文件转换为RKNN模型文件;使用RKNN API编写应用程序,加装RKNN模型文件,进行人脸和口罩检测的推理应用;编译应用程序,并把应用程序和RKNN模型文件部署到RK3399 Pro硬件平台;应用程序通过UVC协议读取USB摄像头的监控图像,使用RKNN调用推理引擎在NPU上分析图像,然后输出检测结果。

[0043] 本实施例中,检测结果可进一步处理:如集成声音提醒等。

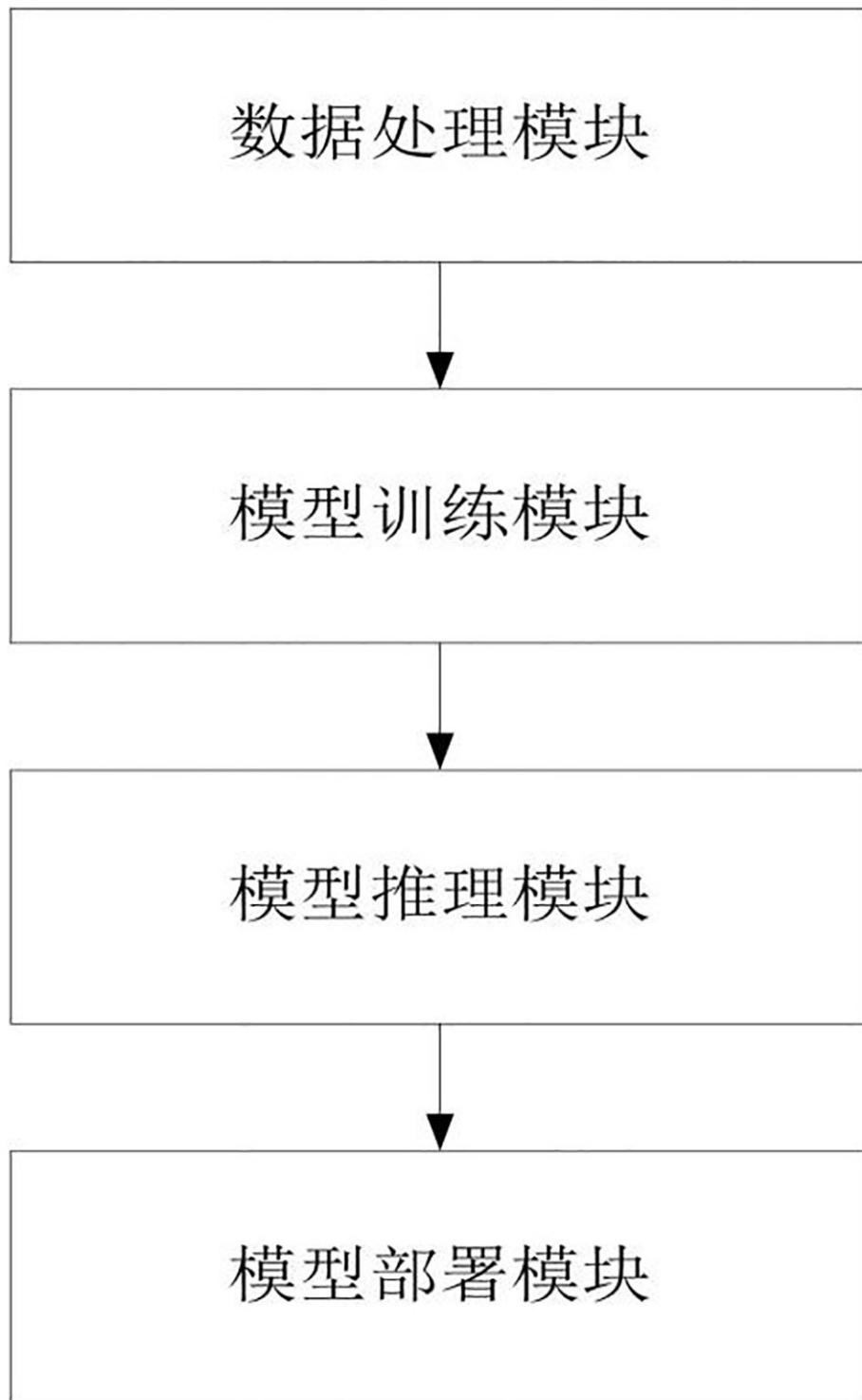


图1

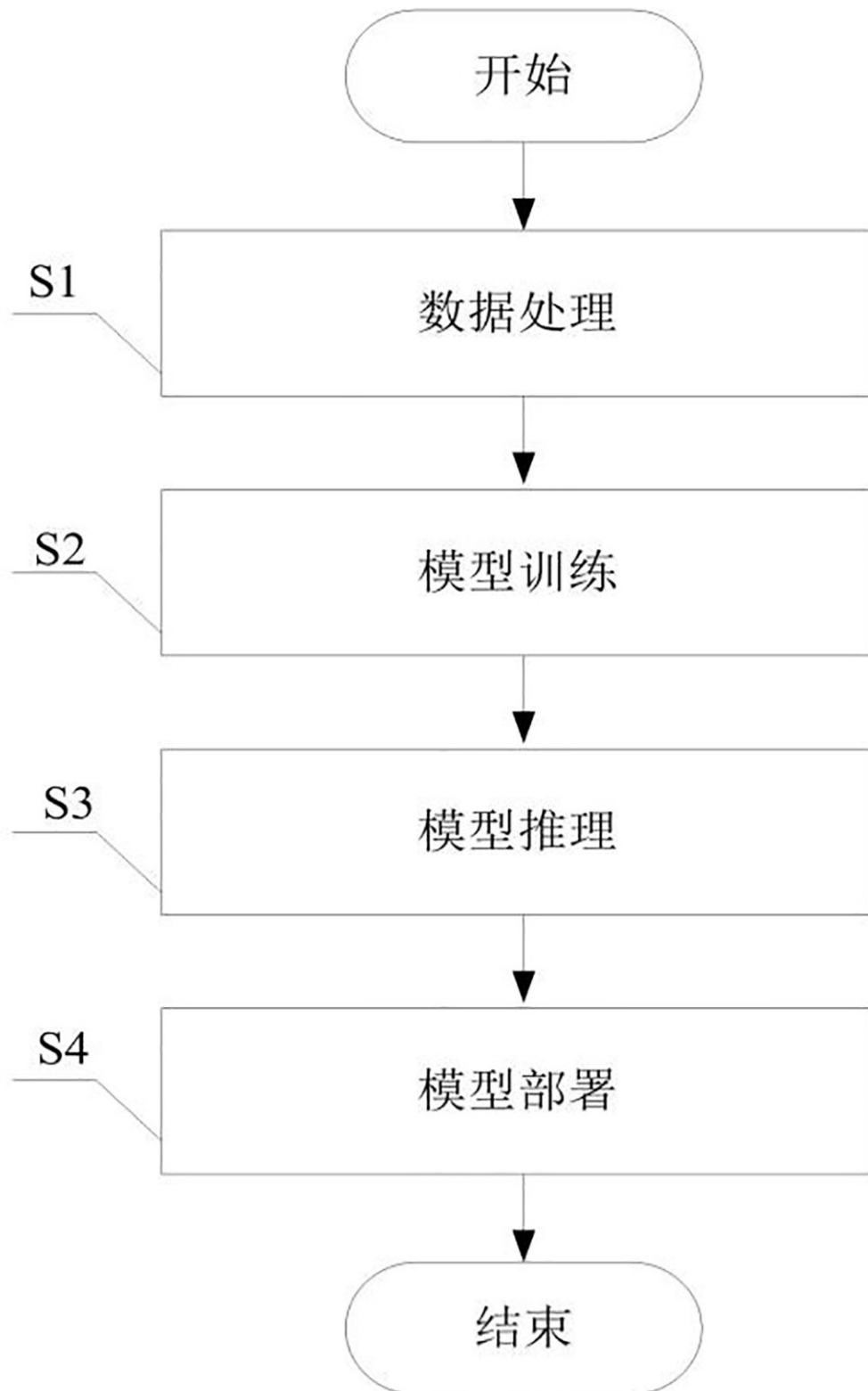


图2



图3