中国人民大学
RENMIN UNIVERSITY OF CHINA

第二十四届“创新杯”

学生课外学术科技作品竞赛

作 品 名 称： 基于华为 Atlas-200 边缘计算开发板的口罩佩戴检测任务

竞 赛 单 元： 常规赛道

参 评 单 位： 高瓴人工智能学院

二零二二年 三 月 二十六 日

基于华为 Atlas-200 边缘计算开发板的口罩佩戴检测任务

刘学旻　于兆洋　卢镜先

（中国人民大学高瓴人工智能学院, 北京市 100872）

摘　要　口罩检测是疫情防控工作中的重要任务。本项目对华为云官方样例进行改进和功能延伸，基于华为 Atlas-200DK 开发者板实现了基本的口罩检测装置，用 PC 机代替华为开发板完成数据采集，通过 scp 指令将数据传输到开发板上，运用 YOLOv3-ResNet18 目标检测算法完成口罩检测任务，并生成具有标注信息的检测结果。经过测试，precision 达到 0.8，recall 达到 0.83，检测效果良好，验证了华为 Atlas-200DK 开发者板在实际场景下的应用潜力。

关键词　口罩检测；华为 Atlas-200DK；YOLOv3-ResNet18；Precision；Recall

Mask Detection Task Based on Huawei Atlas-200

Xumin Liu　Zhaoyang Yu　Jingxian Lu

(Gaoling School of AI, Renmin Univ. 100872, China)

Abstract

Mask detection is an important task in Covid-19 pandemic prevention and control. This project improves and extends the functionality of the official Huawei Cloud example. With Huawei Atlas-200DK, we implemented a basic mask detection device. A PC was used instead of the Huawei development board to complete data acquisition. We transfers data to the development board through scp commands, and uses the YOLOv3-ResNet18 target detection algorithm to complete the mask detection task. Our model generates the labeled information with annotation information. On the test set, the precision reaches 0.8 and recall reaches 0.83, which implies a good detection effect and verifies the application potential of Huawei Atlas-200DK in practical scenarios.

Keywords Mask Detection; Huawei Atlas-200DK; YOLOv3-ResNet18; Precision; Recall

1 引言

近年来，全球肆虐的新冠疫情对人们的生活产生了深刻的影响。目前全球疫情仍不稳定，且出现常态化的趋势，这对疫情防控工作提出了更高的要求。检查人们出入公共场所时的口罩佩戴情况，是日常防疫工作中的重要部分。但人工检测容易出现漏检、错检的情况，也是对人力资源的一种浪费，通过计算机视觉实现口罩识别，能够节省人力资源，提高检测准确性，具有显著的现实意义。

按照应用场景，口罩检测任务可分为主动场景和被动场景两类：主动检测是在地铁站、飞机场等的卡口、闸机处进行单人检测，摄像头获取的图像一般质量较高，但对检测准确性的要求很高；被动场景是在商场、教学楼等公共场合对人流进行检测，存在数据质量低、噪声干扰大的问题。

口罩检测任务主要是基于图像分析与处理技术，对人脸目标进行口罩佩戴与否的检测。检测步骤包括图像获取及预处理，特征提取和有监督二分类。与常规的检测任务相比，口罩检测存在人脸的口鼻以及脸的大半被口罩遮挡，模型难以学到有用的特征，且难以识别不正确的口罩佩戴姿势的问题，因此传统的手工特征提取难以满足应用要求。相比而言，深度网络可以捕获更深层次的复杂特征，能够显著提升检测的速度和准确率，因此被广泛应用于特征提取任务。纵览当前的相关研究工作，口罩检测使用的算法主要有基于深度网络和基于目标检测的算法两类。

基于深度网络的算法主要用于主动场景。按照分类任务实现方法的不同又分为 DNN 网络结合机器学习分类算法和端到端的 DNN 模型两类。其中，结合机器学习分类算法的 DNN 网络算法采用 VGGNet、GoogleNet 网络、ResNet 等网络完成特征提取，再用支持向量机、K 近邻算法等完成分类。这类方法需要通过两个不同的网络分别进行特征提取和分类，存在建模耗时较长，推理速度慢，训练数据较少时机器学习方法容易产生过拟合等缺陷。为此，也有一些研究尝试通过基于 ResNet、DenseNet 等网络的端到端 DNN 模型直接完成检测任务。从有关研究成果来看，检测速度相比于结合机器学习分类算法的 DNN 网络有很大提升。然而，基于深度网络的算法普遍存在模型结构复杂，运行占用大量内存，在计算资源有限的情况下很难部署的问题，其实用性有待商榷。

基于目标检测的算法在一定程度上解决了基于深度网络的算法存在的问题，有成为口罩检测任务主流算法的趋势。目标检测算法分为二阶段检测和一阶段检测两类，其中二阶段检测首先生成可能包含物体的候选区域，再对该区域做进一步分类校准得到最终结果，代表网络有 RCNN 系列，精度高但速度较慢；一阶段算法则基于回归思想，使用一阶网络直接对输入图像进行定位和分类，速度极快，符合实时检测的需要，但精度相比于二阶段法有所下降，其代表网络有 YOLO 系列网络和 SSD 等。其中，YOLOv3 网络通过 ResNet 系列或 Darknet 系列网络实现多尺度特征提取，从而有效捕捉不同尺度的检测目标，能满足现实场景下的检测任务需求，在口罩检测任务中得到了广泛应用。近期研究主要通过改进原始 YOLOv3 网络的结构，在保证精度的基础上尽量减少模型中卷积层的层数，实现模型的轻量化，以在现实场景中得到更好的应用。

总之，作为特殊应用场景下的物体检测算法，口罩检测任务经历了从手工标注特征的传统算法到深度学习算法的演变，主要包括基于深度学习和基于目标检测的算法两类。作为防疫工作的重要技术支持，口罩检测任务除了保证高精度和实时性，更需要在现实场景中得到应用。如何将轻量级、成本低的模型高效部署在硬件上，在硬件限制下最大限度地发挥模型的优势，同样是口罩检测研究的重点所在。

开发板是一种可以进行嵌入式编程的电路板，具有高性能和易部署的优势，是进行模型落地研究的理想平台。在同类产品中，树莓派开发板具有较为成熟的系统，目前已经出现很多依托树莓派开发板进行的研究成果。在口罩检测领域，张丽艳团队的研究将 SSD 网络模型部署到树莓派开发板，并结合舵机和超声波传感器实现具有口罩检测功能的闸机口模型 [1]，具有很好的借鉴价值。

华为 Atlas 200 边缘计算开发板（以下简称华为开发板）是华为公司自研的高性能 AI 应用开发板，搭载昇腾 310 芯片，可以实现模型的快速推理，性能优越。但与同类产品树莓派开发板相比，目前依托华为开发板的研究成果较少，口罩检测领域的研究成果仍然为零*，口罩检测模型仅出现于其官方样例†。我们认为，以华为开发板为平台，在官方样例的基础上实现口罩检测任务的部署和应用，不仅可以在真实场景下检验开发板的实际性能，为未来推进基于华为开发板的实用物体检测技术工作打好基础，更能证明华为开发板在工业化应用上的潜力，为我国自研 AI 产品的大规模推广提供实践支持。

*数据来源于中国知网。截至 2022 年 3 月 26 日，输入关键字“华为开发板”进行检索，暂未找到任何相关的科研工作。

†参见华为云官方 gitee 库样例。

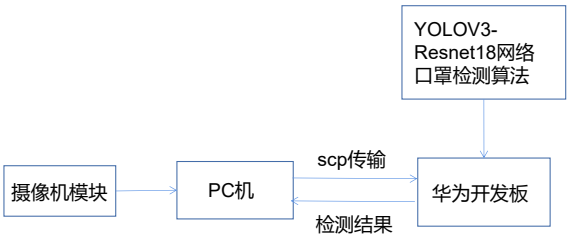

Fig. 1　系统流程图

2　总体设计

本项目构建了基于华为 Atlas 200 边缘计算开发板的口罩检测系统，可以通过摄像头获取实时人脸图像并识别其是否佩戴口罩，可以应用于闸机口等主动场景。项目中以华为开发板作为推理硬件，以 PC 机作为图像采集器和控制模块，系统流程为 PC 端通过图像采集程序，每隔一段时间获取人脸图像，再将其传输到开发板，开发板通过 YOLOV3-ResNet18 算法完成目标检测并计算模型的 precision 和 recall 值，作为评价模型精确度的指标，同时生成含有口罩佩戴检测窗的图像。将其传回 PC 机，可以直观地观察检测结果。‡

本项目是在华为云官方（以下简称官方）gitee 开源代码库中的口罩检测样例的基础上完成的，项目所用的模型采用样例提供的 YOLOv3-ResNet18 网络模型，项目主文件 detect_mask.py 以样例中提供的主文件 mask_detection.py 为基础。

考虑到该样例仅完成了口罩检测任务本身，缺少有效的评价指标，本项目实现了基于 IOU(交并比) 的 precision 和 recall 等评价指标的计算；该样例本身只完成了对单张图片的口罩检测，实用性不强，因此我们增加了通过摄像头实时获取图像输入的功能。出于实际应用考虑，人脸图像输入预期通过开发板的树莓派摄像头接口获取，但限于硬件条件暂时无法实现，因此我们暂将人脸图像获取部分部署在 PC 机上，通过 scp 指令实现开发板和系统的数据传输。在硬件条件允许的情况下，图像采集部分完全可以移植到开发板，实现全自动化的口罩检测。此外，虽然我们没有开发控制闸机口开闭等下游硬件任务，但已经基本实现了关键的上游流程，硬件任务可以视应用场景灵活快速部署。

系统流程图如图1所示。

3　硬件设计

本系统的硬件设计部分主要包括开发板模块和摄像机模块。

3.1　开发板模块

本项目选择华为 Atlas 200 DK 开发者板作为推理硬件，开发板使用 ARM64 架构的 ubuntu18.04 系统作为操作系统，主板型号为 IT21VDMB；以 Micro SD 卡作为其内存硬盘；通过开发板提供的 USB3.0 接口和千兆以太网接口与 PC 机连接，实现数据的传输；采用 Python3.7.5 进行编程。

3.2　摄像头模块

由于硬件条件限制，选择在实验阶段通过 PC 机自带的 720pHD 摄像头完成数据采集工作，以 Python 的 OpenCV 库作为采集并处理图像的主要工具。

4　软件设计

软件设计部分包括开发板软件配置模块和软件工作模块。在进行软件工作前，先对开发板进行软件配置。软件设计部分，本项目通过编写自动拍照脚本，调用 PC 机的摄像头每隔一段时间进行一次拍摄，并将获取到的人像图片存放在系统文件夹下。在开发板与 PC 机通过 USB 线连接的情况下，通过 scp 指令将图像文件上传至开发板工作目录下的文件夹 data/内。运行主程序脚本 mask_detect.py，载入模型，并对 data 文件夹下的图片依次进行推理，进行是否佩戴口罩的检测，计算检测结果的 precision，recall 和 accuracy 值作为模型评价指标，

‡我们的代码以及所用数据集已经开源在Github。

生成带有人脸位置和口罩位置检测框的图片，存放在工作目录下的/out 文件夹中。由于开发板本身不具有图像显示功能，可以通过 scp 指令将/out 文件夹下的内容传输至 PC 机的系统目录下，观察检测效果。

软件设计整体流程如图2所示：

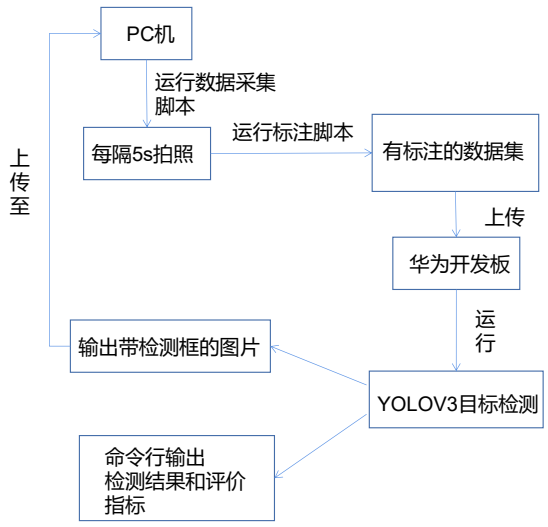

Fig. 2　软件设计整体流程

4.1　开发板软件配置模块

在进行工作前，首先参考华为云官方说明文档完成对开发板的相关软件配置，包括制作 SD 卡，安装 arm64 操作系统和 ubuntu18.04 系统；配置网络连接，使得开发板能够通过共享与之相连的 PC 机的网络或接入路由器上网，以便从华为云 gitee 库下载项目所用的模型；配置 USB 连接，使得在 PC 机上能够通过 ssh 指令登录到开发板，进行数据传输和直接控制。

4.1.1　制作 SD 卡

将开发者板驱动与运行包、ubuntu 操作系统镜像包烧录到 SD 卡，搭建开发板的基本系统架构。实际配置过程参考华为云论坛教程[§]，借助 Etcher 制卡工具将 DD 镜像烧录至开发板，即可完成操作系统搭建和基本的环境配置。

4.1.2　配置 USB 连接

通过 USB 端口将华为开发板与 PC 机连接。华为开发板的 USB 网卡默认 IP 地址为 192.168.1.2，故修改 PC 机的 USB 虚拟网卡地址在 192.168.1.x 网段，并安装 USB 网卡驱动，实现其与开发板的连接，在命令行通过 ssh 指令登录到开发板。

4.1.3　配置网络连接

修改开发板的通信协议，配置其 IP 地址获取方式为 DHCP，eth0 的网关地址与 PC 机中与开发板相连的 NIC 网卡的 IP 地址保持一致。将开发板通过网线与 PC 机相连，即可将 PC 机上的网络共享给开发板。

4.2　软件工作模块

4.2.1　建立数据集

由于本实践项目使用的模型是来自华为云官方的开源口罩检测模型，已经经过训练并具有一定的精确度，因此我们参考张丽艳团队基于树莓派的研究中的数据集设计 [1]，制作了 100 张图片组成的测试集，验证该模型在现实场景下的实际性能，为华为开发板的落地应用提供实践依据。

考虑到现实应用场景，我们选择通过执行图像采集脚本，实时获取人脸图像用于检测。测试集由 100 张在不同背景下、从不同方向拍摄的图片组成，其中 50 张为正确佩戴口罩的图片，50 张为错误佩戴口罩的图片（包括未佩戴口罩和仅用手遮挡面部的情景）[¶]，和现实情景接近，有一定的参考价值。

4.2.2　YOLOv3-ResNet18 目标检测算法

本项目通过主干网络为 YOLOV3-ResNet18 的目标检测算法实现主动场景下的口罩检测。YOLOv3 属于一阶段目标检测算法，相比两阶段目标检测算法速度更快，模型体量也比较小，适合在硬件平台部署。与其他一阶段算法相比，YOLOv3 在输入图像尺寸相当的情况下，比 SSD 检测精度更高，比 Retinanet 检测速度更快，兼顾了检测速度与准确率，是作为口罩检测算法的理想类型。此外，YOLOv3 算法中的 32 倍降采样、16 倍降采样和 8 倍降采样操作符合 Ascend310 芯片适合 16 的整数倍的参数的特点，与其他算法相比能够更好地发挥华为开发板的性能优势[‖]。

在原始 YOLOV3 的基础上，由于 YOLOV3 网络默认的特征提取网络 Darknet53 层数较多，会占用大量计算资源，虽然保证了精度，但模型体

[§]参见华为云论坛帖子。

[¶]为保证各类别的合理性，“正确佩戴口罩的姿势” 和 “错误佩戴口罩的姿势” 来源于应对新型冠状病毒肺炎疫情联防联控机制综合组.《公众和重点职业人群戴口罩指引（2021 年 8 月版）》。

[‖]参见华为云论坛帖子。

量较大，算法复杂，实际部署存在困难，因此该模型选择更加灵活轻便的 ResNet18 残差网络作为特征提取部分。此外，本模型将 YOLOv3 网络的 DBL 中的 Leaky ReLU 替换为更常用的 ReLU，组成 Conv+BatchNorm+ReLU 的组合，该组合基于 Ascend310 处理器进行了专门优化，能更好地提升算法性能。

本系统输入模型的待检图像尺寸为 352*640，将特征图输入到 ResNet18 网络进行特征提取。首先经过一层 7*7 标准卷积，再进行一次最大池化，之后将池化层的输出通过八个二层的残差块。残差块的结构如图3所示，设其输入为 X，weight layer 表示卷积层，包括 conv 卷积层和 batch normalization 批标准化层，ReLU 为激活函数，输入 X 经过变换后与卷积层的输出结果相加。最后通过一个平均池化层。为了更好地检测不同尺度的目标，YOLOv3 网络在残差网络的不同阶段产生三个特征图进行后续的目标检测，最后通过非极大抑制 (NMS) 算法输出最终预测框和预测结果。图4所示为完整的 YOLOv3-ResNet18 结构。

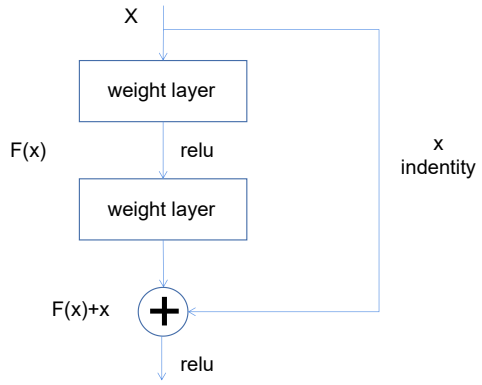

Fig. 3　残差块结构示意图

4.2.3　模型转换工作

为提高模型推理速度，充分发挥华为开发板的推理优势，本项目通过开发板系统自带的 atc 工具进行模型格式转换。华为 Atlas 200 DK 支持通过 ATC(ascned tensor compiler) 工具将 Caffe、Tensorflow 等开源框架的网络模型以及单算子 json 文件转换为🄴腾 AI 处理器支持的.om 格式离线模型，转换过程中可以实现算子调度优化、权值数据重排、内存使用优化等，可以在保证原模型性能的基础上，进一步提升模型推理效率，并有效缓解口罩检测模型相对复杂导致的内存占用量大的问题。

从华为云官方 gitee 开源代码库下载基于 Tensorflow-YOLOv3 的口罩检测模型的原始模型文件 mask_detection.pb，保存在工作目录下的/model 文件夹内，执行 atc 模型转换命令：

```
atc ——input_shape=” images:1,352,640,3” ——
    input_format=NHWC ——output=” ./
    mask_detection” ——soc_version=
    Ascend310 ——framework=3 ——model=” ./
    mask_detection.pb”
```

其中，framework 指定模型框架类型，framework 为 3 即表示模型为 tensorflow 架构；model 指定转化的模型对象，此处为从官方 gitee 库下载的 tensorflow 模型权重文件 mask_detection.pb；input_format 指定模型输入的格式，通常为’ NCHW’ （number，channel，height，weight）或’ NHWC’；input_shape 指定模型输入的尺寸，即每次输入一张图片进行推理，图片尺寸为 352*640；output 指定转换出的.om 文件的路径和名称，通常和原模型相同；soc_version 指定芯片版本，取值固定为 Ascend310。

转换后的.om 文件保存在 model 文件夹下。后续可以通过 python-acllite 库的 AclLiteModel(MODEL_PATH) 函数载入.om 模型，通过 model.execute(INPUT_DATA) 函数完成模型推理。

4.2.4　模型性能测试

Precision 和 recall 指标被广泛应用于模型性能测试，在目标检测算法中同样具有很好的参考价值。其一般的计算公式为：

$$Precision = TP/(TP + FP),$$

$$Recall = TP/(TP + FN).$$

其中 TP 为正例（佩戴口罩且姿势正确）中预测正确的样本数，FP 为正例中预测错误的样本数，FN 为负例（佩戴口罩但姿势不正确，或未佩戴口罩仅用手遮挡）中预测错误的样本数。

但在目标检测算法中，TP、FP 和 FN 的计算方式与其他任务有所不同。由于目标检测通过生成检测框确定检测目标的大概位置，再进行检测任务，检测框的准确程度是衡量模型精确度的重要工具，引入 IOU（交并比）评价指标，其计算公式为：

$$IOU = intersection/union.$$

其中 intersection 为生成的检测框和预先标注

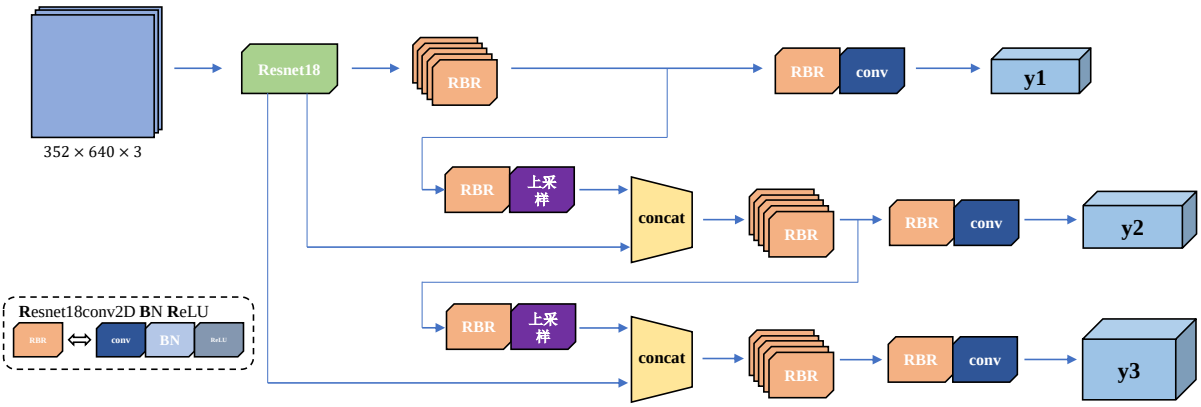

Fig. 4　YOLOv3-ResNet18 网络结构示意图

的检测框的交集，union 为其并集。

相应地，目标检测算法中 TP 和 TN 的定义为：

- $TP = IoU > 0.5$ 的检测框数量，
- $FP = IoU <= 0.5$ 的检测框，或者是检测到同一个 GT 的多余检测框的数量，

按照此定义进行评价指标的计算。

模型对输入的图片进行推理后，会生成口罩检测框，并以列表形式返回其左上角和右下角像素在图片上的坐标。为了计算 IOU 值，基于 OpenCv 库对测试集中正确佩戴口罩的图片进行检测框标注，将检测框左上角和右下角像素点在图片上的坐标按与模型生成的检测框相同的格式，保存为与对应图片同名的.csv 文件，随图片上传到开发板进行推理性能测试。

将测试集图片分为正例和负例两部分分别进行检测。对于正例图片，如果检测结果为“戴口罩”，则根据模型返回的检测框坐标和预先标注的实际坐标计算预测框的 IOU 值。如果 IOU 值大于 0.5，视为检测正确，TP 值加一；反之，FP 值加一；如果检测结果为“未戴口罩”，则 FP 值加一。对于负例图片，如果检测结果为“戴口罩”，FN 值加一；如果检测结果为“未戴口罩”，则 TN 值加一。

根据所得的 TP, FP, TN, FN 值计算 precision 和 recall 值，作为模型性能的评价指标。

5　成果与分析

本项目制作了由手动标注的 120 张图片组成的测试集，并基于华为开发板完成了测试集上的口罩检测任务。图六的检测结果显示，本模型的 precision 值达到 0.8，recall 值达到 0.83，accuracy 值达到 0.82；检测效果较好；从模型输出的带有检测框的图片来看，本模型对于错误佩戴口罩和用手遮挡的姿势有一定的鉴别能力，能实现较为准确的口罩检测。但对于人脸侧面面向摄像头或离摄像头距离较远的情形存在一定的误差，对于与环境颜色相近的口罩识别能力有限，尚有一定的提升空间。

6　结束语

本项目将基于 YOLOv3-ResNet18 网络的口罩检测算法部署在华为 Atlas 200 DK 开发者板上，并在近实际应用场景下实现口罩检测任务，取得了较好的检验结果，验证了模型本身的检测效果和开发板本身的推理性能，证明了华为开发板应用于实际场景任务开发的可行性。当然，由于硬件条件限制，本项目还存在一定的不足。后续会在硬件条件支持的情况下将数据采集算法部署至开发板，实现全自动的实时口罩检测。此外，本项目使用的 YOLOv3-ResNet18 网络尚有改进空间，可以通过将 ResNet 网络替换为 DarkNet 网络等方式进一步提升模型性能，提高检测效率，更好地为疫情防控工作做出贡献。

致　谢　在此，我们谨对提供本项目所用的华为开发板的指导老师鲁蔚征老师，为我们的项目提供无私帮助和宝贵意见的指导教师胡迪老师，对提供本项目所用模型及主函数原型的华为云官方 gitee 库负责者，对华为云论坛上在我们的开发板部署工作中热心提供技术指导的的工作人员，表示由衷的感谢！

参考文献

[1] 张丽艳, 赵艺璇, and 李林. 一种基于树莓派的口罩自动检测装置. 大连交通大学学报, 43(1):4, 2022.

[2] 张琴涛. 基于改进 YOLOv3 的口罩佩戴检测算法研究. PhD thesis, 天津工业大学.

[3] 曾成, 蒋瑜, and 张尹人. 基于改进 yolov3 的口罩佩戴检测方法. 计算机工程与设计, 42(5):8, 2021.

[4] 柯鑫, 张荣芬, and 刘宇红. 嵌入式口罩佩戴检测系统研究与实现. 智能计算机与应用, 11(7):6, 2021.

[5] 王远大. Ucloud 开放人脸口罩检测服务借助 ai 算法加快疫情防控. 通信世界, (5):2, 2020.

[6] 董艳花, 张树美, and 赵俊莉. 基于残差结构的 ssd 口罩检测. 计算机技术与发展, 31(12):6, 2021.