基于 YOLO-v3 的安全帽佩戴检测

刘瑶 周丽

四川大学锦城学院 计算机与软件学院,四川 成都 611731

【摘 要】：本文简要介绍了目标提取网络的发展，YOLO-v1 网络借了谷歌网络的主要网络结构，YOLO900 相比 YOLO-v1 提高了训练图像的分辨率，YOLO-v2 的水平是目前目标检测领域内最先进的，YOLO-v3 在 cocotest-dev 上达到 57.9%。随着国家大力完善社会基础建设，建筑业安全生产的需求也逐步进行完善。随着我国的经济水平不断提高，我国的建筑业事故造成的损失仍然较大。近年建筑行业生产安全事故造成的损失，施工现场由于违章违规造成的安全事故频繁发生。基于这样的背景下，本文讲解了一种基于 YOLO-v3 实现的简便的安全帽佩戴检测系统。

【关键词】：目标检测；YOLO-v3；深度学习

引言

随着国家大力完善社会基础建设，建筑业安全生产的需求也逐步进行完善。随着我国的经济水平不断提高，我国的建筑业事故造成的损失仍然较大。近年建筑行业生产安全事故造成的损失，施工现场由于操作不合规定 造成的安全事故频繁发生。在施工过程中，因为监管力度较弱，工人不佩戴安全帽而引发安全事故频频发生。为了改善这种情况，可以通过监管佩戴安全帽这一规章制度改进，需要对施工人员的安全帽佩戴进行检测，用于减少因为作业人员没有佩戴安全帽而造成的事故伤害。

1 目标检测发展简介

以往的目标检测方法检查准确率低，通过手工设计实现。随着我国经济的发展，深度学习算法获得学者喜爱，很多科研人员提出了一系列基于卷积神经网络的目标检测模型被很多研究人员提出。人工智能领域的研究人员在将传统使用的滑动窗口加手工设计特征使用替换为候选区+卷积神经网络模型，设计了区域类型的深度学习神经网络，在部分数据集上，将目标检测的平均准确度大幅度提升。

通过以上的规律及长时间的研究，科研人员提出了快速区域卷积神经模型和超快区域卷积神经模型被一些人工智能领域的研究员提出了，此网络提高了准确率和增加了检测速度，使得各项指标可以达到较高的水准。

YOLO 检测模型被一些目标提取领域的专家提出。一些图像识别的专家提出了 SSD 目标提取检测模型，该算法在检测精度和检测时间上均取得了较好的效果，在这样的情况下，YOLO 的作者提出了 YOLOv2 和 YOLOv3 目标检测特征提取模型，其中 YOLOv3 的检测效果最好，在一些数据集上在较短的时间达到较好的效果，与 RetinaNet 达到的效果相仿，性能相仿但速度比之快 3.8 倍。目前，YOLOv3 在目标检测领域有较好的特征提取效果。

2 YOLO 网络发展简介

在参考谷歌网络的分类网络结构下，YOLO 的作者提出了 YOLO-v1 网络。但是，YOLO 未使用与谷歌网络的 inceptionmodule，而是使用 1×1 和 3x3 卷积层替代。YOLO-v1 全连接层的输出维度是 S×S×(B×5+C)S×S×(B×5+C)，因为网络在最后使用全连接层进行类别输出。YOLO-v1 网络的速度比 VGG16 的速度快，但是 YOLO-v1 网络的准确率稍差。

YOLO 作者利用 faster-rcnn 的部分结构，基于使用卷积核替代输出层的思想，提出了性能更优的 YOLO-v9000。联合使用物体检测标注数据和物体分类标注数据训练物体检测模型，提高了物体定位准确度和其他相关性能，对部分神经元的设计进行了精进。YOLO9000 在识别种类个数、识别的精准度的百分比、和识别定位精准度性等各个性能方面都有大幅度提升。

YOLO-v2 的速度要快过其他检测系统 FasterR-CNN，ResNet，SSD 等，所以代表着目前目标特征提取领域的最先进水平。它的速度与精确度可以进行权衡。YOLO-9000 可以实时地检测提取种物体的图像特征因为其 WordTree 使用 wordTree 来混合检测各个数据集之中的数据。

通过改变网络模型结构的大小来评估速率与精确度，YOLO-v3 在准确度可以上达到较高水平，得更加复杂化。执行检测任务通过 YOLO-v3 的先验检测系统将分类器或定位器实现。此模型可以应用于图像的多个角度。检测结果通过那些评分较高的区域表现。此外，相对于其它目标检测模型，YOLO-v3 使将图像分割为不同的小块，将一个单一神经元作用于图像，通过这种方法预测每一块的边框和概率大小判断其类型。

YOLO 模型相比于其他传统类型的模型有一定的优势,它检测了图像中的总体信息而不是只提取局部信息，所以在测试时会查看图像的总体信息只需使用单一神经网络进行评估预测,这令 YOLO-v3 非常快，一般它比 R-CNN 和 FastR-CNN 都要快。

3 YOLO-V3 详解

3.1 Darknet-53 与 YOLO-v3

Darknet-53 是 YOLO-v3 的主干网络，darknet-53 由 52 组卷积神经网络层和有一个全连接层组成。其中的卷积层部分可以分为一层普通卷积层和五个重复的残差元。在网络中普通卷积层包括：1 个 2 维度卷积,；1 个批正则化层，将输入数据 x 进行正则化；1 个斜率是 0.1 的 LeakyReLU 层，LeakyReLU 是在 ReLU 基础改进后的激活函数。5 组残差单元分别由 1 个单独的卷积层与一组重复执行的卷积层构成，重复执行的卷积层分别重复 1 次、2 次、8 次、8 次、4 次。在每个重复执行的卷积层中，先执行 1x1 的卷积操作，再执行 3x3 的卷积操作，过滤器数量先减半，再每组残差元的第一个单独的卷积层操作均是一次步长为 2 的卷积操作。

darknet-53 网络只有前 52 层的卷积网络被 YOLO-v3 使用了。YOLO-v3 是一个全卷积层的神经网络。YOLO-v3 进行了三次下采样。按照 2 的 5 次方，2 的 4 次方和 2 的三次方的顺序。通过上采样可以使得网络越深则特征表达效果越好，如果直接使用第四次下采样的特征进行检测，进行 2 的 4 次方倍降采样检测，就只能提取到浅层特征，导致效果不好。如果把在 2 的 5 次方倍降采样得到的特征图的图像尺寸大小增大，变成 2 的 4 次方倍采样后的维度，就可以得到 2 的 5 次方倍降采样后的图像特征。通过这样的途径可以实现网络深层化的功能。

表 1YOLO-v3 网络结构

Amount	TYPE	FILTERS	SIZE	OUTPUT
1	conv	32	3×3	256×256
1	conv	64	3×3/2	128×128
1	conv	32	1×1	
1	conv	64	3×3	
1	residual			128×128
1	conv	128	3×3/2	64×64
2	conv	64	1×1	
2	conv	128	3×3	
2	residual			64×64
1	conv	256	3×3/2	32×32
8	conv	128	1×1	
8	conv	256	3×3	
8	residual			32×32
1	conv	512	3×3/2	16×16
8	conv	256	1×1	
8	conv	512	3×3	
8	residual			16×16
1	conv	1024	3×3/2	8×8
4	conv	512	1×1	
4	conv	1024	3×3	
4	residual			8×8

3.2 特征提取原理

对于 YOLO-v3 关于边界框的初始图像大小还是采用 YOLO-v2 中的 K 均值聚类的算法实现的这种先验知识对于边界框的初始化的益处还是很明显的，毕竟过多的边界框对于算法速度的影响还是非常显著的，虽然效果非常显著的。YOLO-v2 借鉴了 faster R-CNN 的部分机制，不同的是，采用 K 均值聚类的方法来确定默认框的尺寸。Redmon 修改了 K 均值聚类算法中关于距离的定义，使用的是 IOU 距离。YOLO-v3 选择的默认框有 9 个。其尺寸可以通过 K 均值聚类算法在数据集上聚类得到。

YOLO-V2 只适用于一个提取目标只属于一个类的情况，但是在生活中情况总是复杂的，在此基础上，YOLO-V3 提出更为复杂和合理的模型，将单标签的情况改善为多标签识别的情况。但在一些复杂多因素影响的情况下，一个目标可能属于多个类，所以 YOLO-v3 将 Softmax 用多个单独的逻辑分类器替代解决多标签分类问题，而且准确度不会下降。

4 系统结构

4.1 数据集介绍

数据集总共有 7581 张图像，包含 9044 个佩戴安全帽的 bounding-box，以及 111514 个未佩戴安全帽的 bounding-box，所有的图像用 labelimg 标注出目标区域及类别。其中“hat”表示佩戴安全帽，“person”表示普通未佩戴的行人头部区域的 bounding-box。

4.2 训练模型

下载 github 上开源的 YOLO-v3 的预训练模型进行加载。然后加载数据集，并对数据集进行处理，压缩送入预训练模型进行训练，通过加载训练好的模型参数，对每一帧中的人进行预测，判别是否佩戴了安全帽。

4.3 系统设计

通过摄像头对工地的员工工作状态进行实时检测，通过对摄像头传回的页帧进行处理，通过目标提取网络对特征进行提取，识别。判别员工是否正确佩戴安全帽，如出现员工未佩戴安全帽或佩戴不正确的情况则通过控制硬件设施对此员工进行提醒，记录。通过基于 YOLO-v3 的预训练模型的驯练，得出适用于检测安全帽佩戴的安全帽检测模型，完成对具体场景的监测。在数据库的设计上，由于一个工地的具体需要记录的人脸及名字信息较少，可直接采用轻量级的数据库 sqlite3 进行设计，人机交互界面可采用 qt 进行设计，依次将各个模块有序链接。

图 1 结果展示图

5 结语

此系统操作简单消耗成本较低，识别准确度和效率较高，可以作为建筑工地安全保障的辅助工具。此系统可根据需求改变相应结构，从而应用于其他的工作场景中。

参考文献：

[1] 鞠默然,罗海波,王仲博,何淼,常铮,惠斌.改进的 YOLOV3 算法及其在小目标检测中的应用[J].光学学报,2019,39(07):253-260.

[2] 施辉,陈先桥,杨英.改进 YOLOv3 的安全帽佩戴检测方法[J].计算机工程与应用,2019,55(11):213-220.

[3] 赵华峰. 基于单片机的家居智能系统设计与实现[J]. 信息与电脑(理论版),2018(22):134-136.

[4] 何波.基于单片机的人工智能语音识别机器人的设计分析[J].电子世界,2018(08):72+74.

[5] 蔡自兴.中国人工智能 40 年[J].科技导报,2016,34(15):12-32.