

目标检测算法用于安全帽佩戴检测

周 丽

成都锦城学院 计算机与软件学院 四川成都 611731

摘要: 在复杂工地场景中, 为保证施工人员的人身安全, 常常要求佩戴安全帽。以往靠人工监督的方法, 不仅消耗人力资源, 且易出现漏检、误检等情况。因此, 本文采用深度学习的YOLO目标检测算法, 基于网络公开安全帽数据集, 对算法的预训练模型进行微调, 分两阶段迁移学习。最终由实验结果可知, 该方法能得到的准确度更高的模型, 进而实现工地场景中的安全帽佩戴检测。

关键词: 目标检测; YOLO; 安全帽检测; 迁移学习

引言:

在施工工地, 为保证施工人员的安全, 最基本的一条要求就是正确佩戴安全帽, 安全帽可以有效避免和减小事故带来的伤害。传统的人工巡检监控安全帽佩戴情况耗时耗力, 容易发生误检。利用机器视觉和深度学习技术替代人力, 来对施工工地人员佩戴安全帽进行检测识别, 效率更快, 准确度更高。

传统的安全帽检测多是基于图像处理算法, 如先定位人头位置区域, 再通过训练特征分类器来实现安全帽的检测。这种方法易受环境影响, 如雨天, 遮挡等情况则检测效果不佳, 而施工工地又复杂多变, 所以传统特征提取的方法效果并不理想, 受限较多。

随着深度学习的快速发展, 基于深度学习的目标检测算法能够自动提取图像中不同区域的特征^[1], 通过学习样本中的有效特征, 从而建立起泛化能力较强的目标检测模型。因此, 本文基于已在COCO数据集上训练好的YOLOv3目标检测模型, 通过迁移学习安全帽佩戴样本图像数据, 最终训练得到能实现工地环境中的安全帽实时检测模型。

一、YOLO目标检测算法

基于卷积神经网络实现目标检测, 需要解决目标的定位问题和识别问题, 即模型需要输出目标的位置和类别。通常目标的位置可以由x、y、w、h四个值确定, 即目标外部矩形框左上角顶点的坐标和矩形框的宽、高值。而目标类别则是预先设定好的范围内的取值, 例如COCO数据集有80个生活场景中常见的物体类别, 包括人、车、动物、水果、家具等。

常见的目标检测算法分为两类, 基于建议框的方法和免建议框的方法。基于建议框的方法又称为两阶段方法, 即将目标检测分成建议框生成和建议框分类、回归两个阶段, 如Faster-RCNN和Mask-RCNN都是基于建议框的方法; 而免建议框方法又称为单阶段方法, 即没有建议框搜索的步骤, 如YOLO和SSD等属于单阶段算法。通常两阶段的目标检测算法精度高, 但检测速度不如单阶段的算法快。

YOLO方法的基本思想是, 先将全图划分成 $N \times N$ 的格子, 每个格子负责对落入其中的目标进行检测, 一次性预测所有格子所含目标的边界框以及所属类别概率。

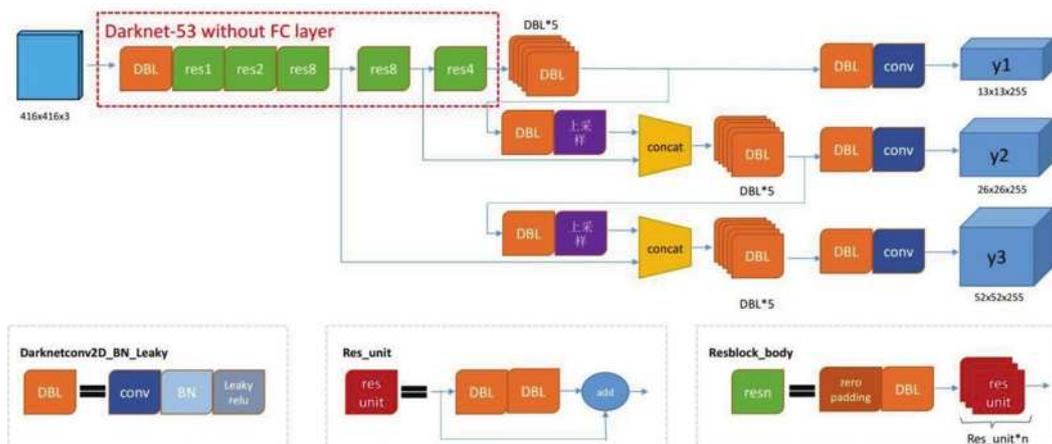


图1 YOLOv3模型结构图

YOLOv3 在每个格子里会预测 3 个锚定框，训练时会选取其中和 GT（即正确答案的边框）的 IOU（最大交并比）值最大的边框作为预测结果。（见图 1）

YOLOv3 先使用 Darknet53 结构来实现特征提取，再通过 FPN（Feature Pyramid Networks）网络层来实现多尺度目标的特征提取。其中，Darknet53 是全卷积结构，包含了 53 个卷积层，主要由 3x3 和 1x1 滤波器组成，同时加入了残差结构，以减少因网络层次过多带来的梯度消失问题。Darknet53 去掉了池化层和全连接层，特征图尺寸变换则是通过改变卷积核的步长来实现。FPN 网络层采用类似特征金字塔的方式，通过上采样和特征融合，最终输出三种不同尺寸的检测结果。YOLOv3 在 COCO 数据集上的准确度与 SSD 相当，但速度却是 SSD 的三倍，YOLOv3 相比 YOLOv1 和 YOLOv2 模型，在小目标检测上性能更突出。

二、安全帽检测与结果分析

1. 安全帽数据集

本文采用的安全帽数据集来自网络公开数据集，共有 6057 张日常生活图片，场景包含工地、大学监控视频和普通场景。为提高模型的准确度和泛化能力，同时考虑到正负样本的均衡，会先在正负样本数据里均匀采样，其次对样本数据进行水平翻转、缩放等操作对数据进行数据的扩充。



图 2 安全帽数据集样本图片

2. 先验框尺寸聚类

YOLOv3 预训练模型中，使用的 COCO 数据集图像会被缩放到 416x416 大小，图像中包含的目标大小不固定，既有较大的目标，也有小型目标，所以先验框的大小分成大、中、小三类，每类又有三种尺寸，分别是属于小目标尺寸的 (10x13)、(16x30)、(33x23)，中等目标的尺寸 (30x61)、(62x45)、(59x119)，以及大目标尺寸的 (116x90)、(156x198)、(373x326)。因为本文采用的安全帽数据中，目标仅为安全帽，所以目标的宽高比和尺寸大小变化不大，若继续采用基于 COCO 数据的先验框大小，则会影响训练进度。因此，在开始训练新的数据前，需要先使用 K 均值聚类算法，对训练样本中

所有目标的宽和高进行 9 个类别的聚类，最后得到 9 个值 (11x13)、(13x17)、(16x17)、(18x21)、(22x26)、(30x35)、(47x54)、(88x103)、(204x236)，作为 YOLOv3 模型训练时的先验框大小。

3. 实验结果

为达到更好的训练效果，训练时分两个阶段进行。第一阶段除了最后三个输出层参与训练外，其余层均被冻结，学习率采用较小值 0.001，训练样本批大小为 12，训练 50 轮。第二阶段放开所有层，包括特征提取层和输出层，进行微调，学习率采用更小值 0.0001，训练样本批大小仍采用 12，在前一阶段基础上继续训练 50 轮。从图 3 可以看出，经过第二阶段的训练，训练样本损失函数数值能得到进一步的降低。

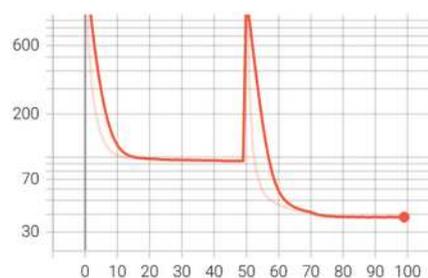


图 3 训练集损失值曲线图

三、总结

本文采用迁移学习的方法，使用 YOLOv3 模型来实现工地场景中的安全帽佩戴检测，以解决安全帽检测实时性和准确性等问题。训练模型时，本文基于 COCO 数据集上的预训练模型，在网络公开安全帽数据集上进行两阶段迁移学习，第二阶段在第一阶段基础上对所有层进行微调，从而得到了更高精度的安全帽检测模型。由实验结果可知，该模型可以实现工地场景中的实时安全帽佩戴检测。

参考文献：

- [1] 李天宇, 李栋, 陈明举, 吴浩, 刘益岑. 一种高精度的卷积神经网络安全帽检测方法[J]. 液晶与显示, 2021, 36 (07): 1018-1026.
- [2] 张勇, 吴孔平, 高凯, 杨旭. 基于改进型 YOLOV3 安全帽检测方法的研究[J]. 计算机仿真, 2021, 38 (05): 413-417.
- [3] 王成龙, 赵倩, 郭彤. 基于面部特征的深度学习安全帽检测[J]. 上海电力大学学报, 2021, 37 (03): 303-307.