

基于深度学习的办公室吸烟行为检测

魏袁慧, 吴四九, 刘天锴, 谭熙

成都信息工程大学计算机学院, 四川 成都

收稿日期: 2023年3月16日; 录用日期: 2023年5月4日; 发布日期: 2023年5月16日

摘要

吸烟有害健康, 为优化办公环境以及保证办公室人员身心健康。随着深度学习中卷积神经网络(Convolutional Neural Network, CNN)在目标检测领域上的发展, 其目标检测方法有: 单阶段检测(YOLO、SSD、RetinaNet等)、双阶段检测(Fast RCNN、Faster RCNN、Cascade RCNN等)。相比于传统的手工设计特征算法, 基于深度学习的方法通过学习大量标注数据来自行进行特征的学习和提取, 并预测或识别出结果, 基于深度神经网络的目标检测方法具有更好的特征提取能力和分类识别效果。本文设计了采用YOLO深度学习算法的办公室吸烟行为检测方法。通过网络公开数据集收集的吸烟数据集, 经过对数据集的整合与调整形成最终进行实验的吸烟行为检测数据集。用吸烟行为检测数据集分别训练YOLOv5, YOLOv6, YOLOv7, YOLOx四个模型, 通过对比训练产生的结果得到办公室吸烟行为检测的最佳训练模型。实验结果表明, 在办公室吸烟行为检测实验中, YOLOv5为检测效果最优异的模型, 其精确度均值(mAP): 76.6%, 平均推理时间: 17.1 ms。

关键词

深度学习, 吸烟检测, YOLO模型

Office Smoking Behavior Detection Based on Deep Learning

Yuanhui Wei, Sijiu Wu, Tiankai Liu, Xi Tan

School of Computer Science, Chengdu University of Information Technology, Chengdu Sichuan

Received: Mar. 16th, 2023; accepted: May 4th, 2023; published: May 16th, 2023

Abstract

Smoking is harmful to health, in order to optimize the office environment and ensure the physical and mental health of office staff. With the development of Convolutional Neural Network (CNN) in

the field of target detection, the methods of Convolutional Neural Network detection include one-stage detection (Yolo, SSD, RetinaNet, etc.) and two-stage detection (Fast RCNN, Faster RCNN, Cascade RCNN, etc.). Compared with the traditional hand-designed feature algorithm, the method based on deep learning can learn and extract feature by learning a lot of labeled data, and predict or recognize the result, the object detection method based on deep neural network has better feature extraction ability and classification recognition effect. In this paper, Yolo deep learning algorithm is used to detect office smoking behavior. The smoking data set collected through the web-based open data set was integrated and adjusted to form the smoking behavior detection data set for the final experiment. Four models, Yolov5, YOLOV6, Yolov7 and YoloX, were trained with smoking behavior detection data set. The best training model of office smoking behavior detection was obtained by comparing the results of training. The results showed that YOLOV5 was the best model in the detection of office smoking behavior. The mean accuracy (mAP) was 76.6% and the mean reasoning time was 17.1 ms.

Keywords

Deep Learning, Smoking Testing, YOLO Model

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

办公室是公共场所，根据卫生部《公共场所卫生管理条例实施细则》第十八条：室内公共场所禁止吸烟。所以在办公室禁止吸烟不仅是国家的明令要求，也是创建文明、和谐、健康社会与生活的必然要求。但就目前公司禁烟情况来看，还任重道远。

国内外还未将深度学习应用于办公室吸烟行为检测场景，本文首次作出尝试，将深度学习中的卷积神经网络用于办公室吸烟行为检测。

通过对学者们在多场景下的吸烟行为检测研究的科研课题的学习，本文归类出吸烟检测方法主要包括烟雾检测[1] [2] [3]、手势动作检测[4] [5]和香烟目标检测[6]，其中，由于香烟烟雾形态复杂、透明度高，关键检测因素采集难度大导致检测精确度低；由于吸烟时手势动作与喝水吃饭等动作相似，靠手势动作判断吸烟的误检率高。因此本文采用香烟目标检测的方法进行吸烟行为识别，将办公室图像信息输入吸烟检测算法中，对香烟目标进行检测。

基于深度学习的目标检测方式，对光线、形变等不确定变量的鲁棒性强，且当今基于深度学习的目标检测，检测的精度和效率得到了很大提升，目标检测的YOLO系列算法已经迭代多代。故本文采用基于YOLO [7] [8] [9] [10]模型的目标检测方法，用吸烟行为检测数据集分别训练YOLOv5, YOLOv6, YOLOv7, YOLOx四个模型，通过对比训练产生的结果得到办公室吸烟行为检测的最佳训练模型。

2. 模型检测原理

本文实验所用四种模型均为单阶段目标检测模型，即对目标直接分出类别，并对边框进行回归，检测原理如图1所示，相比于双阶段目标检测[11] [12]模型，单阶段目标检测模型省去了区域提名(Region proposal)。

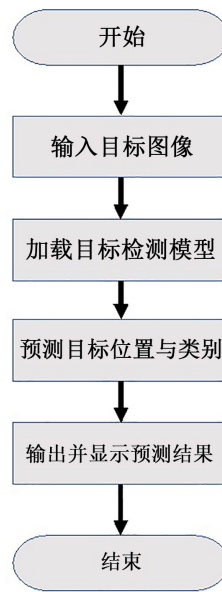


Figure 1. Detection flowchart
图 1. 检测原理图

YOLOv5、YOLOv6、YOLOv7 和 YOLOx 四个模型出自于不同的团队，网络结构各有不同。如图 2 所示，为 YOLOv5 的网络模型。YOLOv5 的骨干网络为 CSPDarknet53 (含 SPPF)，Neck 为 FPN + PAN 结构。YOLOv6、YOLOx 和 YOLOv7 的检测思路和 YOLOv5 相似。其中 YOLOx 的主要改进是引入了 Decoupled head 解耦头的结构，如图 3 所示。YOLOv6 骨干网络由 YOLOv5 的 CSPDarknet 换为了 EfficientRep，neck 基于 Rep 和 PAN 构建了 Rep-PAN。YOLOv7 在 backbone 加入 E-ELAN，E-ELAN 对基数(Cardinality)做了扩展(Expand)、乱序(Shuffle)、合并(Merge Cardinality)，能在不破坏原始梯度路径的情况下，提高网络的学习能力。

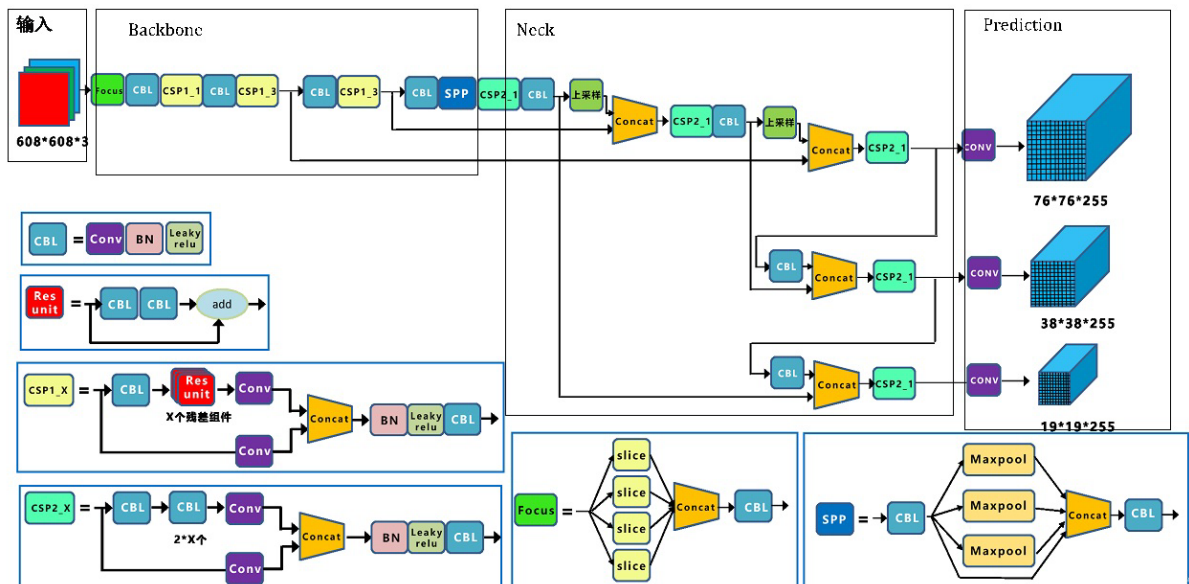


Figure 2. Structure diagram of YOLOv5
图 2. YOLOv5 网络结构图

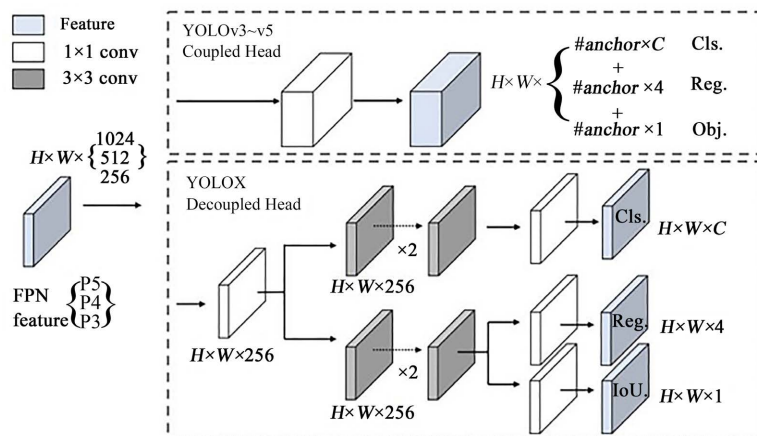


Figure 3. The structure of the YOLOX decoupled head
图 3. YOLOx 解耦头的结构

3. 数据集与实验配置

3.1. 数据集采集及制作

通过网络公开数据集收集吸烟检测数据集，数据集中包含图片信息与标签信息。由于收集的数据集中有部分数据损坏或不满足办公室吸烟检测条件，所以对收集到的数据集进行重新整合与修改，并添加自己通过 Label Img 制作的数据，最终形成本文实验所用包含 4857 张图片的办公室吸烟检测数据集。由于不同模型训练的数据集格式不同，所以，该数据集被制作作为 YOLO 格式和 VOC 格式两种格式进行实验。

3.2. 配置

本次实验在 AMD Ryzen 5 5600H with Radeon Graphics 处理器，主频为 3.30 GHz，内存为 16 Gi B，显卡型号为 NVIDIA GeForce RTX 3050 的 PC 机上进行，在 CUDA11.8 环境下利用深度学习框架 Pytorch 对四个模型进行训练与测试，具体实验配置见表 1。训练模型的参数设置如表 2 所示，模型的训练次数 Epoch 设为 100；批次大小 Batch size 是指训练时一次性输入网络的图片数目，与显卡大小和模型大小有关，所以，将 YOLOv5, YOLOv6, YOLOv7 三个模型的 Batch size 设置为 4，将 YOLOx 模型的 Batchsize 设置为 2；输入分辨率均采用模型原始初始值 640，四个模型均加载官方预训练权重进行训练。

Table 1. Experiment environment configuration

表 1. 实验环境配置

内容	配置
处理器	AMD Ryzen 5 5600H with Radeon Graphics
环境	CUDA 11.8
编程语言	Python 3.9
GPU	NVIDIA GeForce RTX 3050

Table 2. Training parameter setting

表 2. 训练参数设置

名称	参数值
epoch	100
Batchsize	4(YOLOv5,YOLOv6,YOLOv7),2(YOLOx)
Img-size	640

4. 结果与分析

4.1. 模型性能对比

由于 YOLOv5, YOLOv6, YOLOv7, YOLOx 四种预训练模型的网络有所差异, 所以在不同的数据集上表现出来的效果都有所不同。为了得到更适用于办公室吸烟检测数据集所对应的权重数据, 用数据集分别对这四种模型进行训练。然后用测试模型加载训练数据集得出的权重, 用含 300 张图片的测试集对四种模型进行性能对比。通过计算四种模型在测试集上的精确率 P 均值(mAP)以及召回率 R 来评价模型的性能, 如公式(1)所示为精确率的计算, 如公式(2)所示为召回率的计算。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

公式(1)和公式(2)中, TP 表示将正类别预测为正确类别的个数; FP 表示将负类别预测为正确类别的个数; FN 表示将正类别预测为负类别的个数。YOLOv5, YOLOv6, YOLOv7, YOLOx 四种预训练模型测试的结果如表 3 所示。



Figure 4. Model comparison
图 4. 模型对比

Table 3. Performance comparison of different models

表 3. 不同模型性能对比

网络类型	测试数量/张	平均检测时间/ms	精确率均值(mAP)/%	召回率/%
YOLOv5	300	17.1	76.6	80.1
YOLOv6	300	27.53	76	64.4
YOLOx	300	35.16	73.1	85.8
YOLOv7	300	32.4	59.7	57.6

由表 3 可知, YOLOv5 模型的 mAP 值最高, 相对于 YOLOv6 模型, YOLOx 模型, 和 YOLOv7 模型分别高出 0.6, 3.5, 和 16.9 个百分点。YOLOx 模型的召回率最高为 85.8%, 其次是 YOLOv5 模型, 为 80.1%, 相比于 YOLOv6 和 YOLOv7 模型分别高出 15.7 和 22.5 个百分点。YOLOv5 的平均检测时间最快, 为 17.1ms。四个模型对吸烟行为检测样例如图 4 所示。

从图 4 的检测对比可以看出, YOLOv5 的预测精确度较高为 0.87, 优于其他三种检测模型。

4.2. YOLOv5 模型训练过程分析

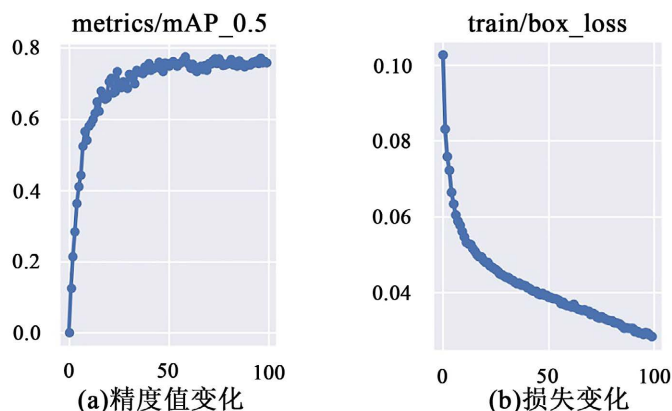


Figure 5. Loss and accuracy of YOLOv5 network

图 5. YOLOv5 实验损失和精度值

综合上述四种模型的测试性能以及测试效果, 本文选用 YOLOv5 模型训练得到的最佳权重数据。图 5 描绘了模型训练时的精确度均值(mAP)和训练损失随迭代次数的变化值。从训练次数 Epochs 的值为 0 开始, mAP 从 0 开始快速上升, 训练损失值由 0.10 开始迅速下降, 标志着网络模型开始对图片特征进行学习。当训练次数 Epochs 为 40 附近时, 精度升至 0.74, 曲线趋于稳定, 波动减少; 损失值降至 0.04, 下降速度减缓。当训练次数 Epochs 为 80 附近时, 精确率曲线趋向为 0.766 的稳定值, 训练损失曲线趋向为 0.028 的稳定值。由此, YOLOv5 模型训练过程稳定, 性能优良。

5. 总结与展望

本文首先对收集的数据进行预处理, 得出更符合办公室吸烟行为检测的数据集, 通过对比 YOLOv5, YOLOv6, YOLOv7, YOLOx 四种预训练模型所得的性能结果和检测效果, 得到最优权重数据。试验结果表明, 相较于其他三种模型, 使用 YOLOv5 模型检测办公室吸烟行为的方法可以快速、精确地识别出吸烟行为, 为办公室禁烟提供了一定保障。

作为国内外基于深度学习的办公室场景下的吸烟行为检测的首次尝试, YOLOv5 的表现并不俗, 我后续将对 YOLOv5 网络模型做出进一步调整, 期待更好的检测结果。

参考文献

- [1] 汪祖云, 廖惠敏, 张日东, 刘鹏宇, 贾克斌. 结合烟雾多特征的出租车司机吸烟行为检测[C]//中国高科技产业化研究会智能信息处理产业化分会. 第十二届全国信号和智能信息处理与应用学术会议论文集. 2018: 368-373.
- [2] 苏翔宇. 基于烟雾多特征的吸烟行为识别算法研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2014.
- [3] 丁宏杰. 基于视频烟雾的吸烟行为识别与研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2013.
- [4] 邸昱卿, 张云伟. 基于人体关键点的吸烟行为识别方法研究[J]. 电视技术, 2022, 46(5): 12-16+29. <https://doi.org/10.16280/j.videoe.2022.05.003>
- [5] 王超. 针对吸烟行为的手势识别算法研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2013.
- [6] 李倩. 基于深度学习的烟支检测技术研究与应用[D]: [硕士学位论文]. 西安: 西安邮电大学, 2020. <https://doi.org/10.27712/d.cnki.gxayd.2020.000027>
- [7] Redmon, J. and Farhadi, A. (2016) YOLO9000: Better, Faster, Stronger. 2017 *IEEE Conference on Computer Vision*

-
- and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21-26 July 2017, 6517-6525.
<https://doi.org/10.1109/CVPR.2017.690>
- [8] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. <https://arxiv.org/abs/1804.02767>
- [9] Zhu, X., Lyu, S., Wang, X. and Zhao, Q. (2021) TPH-YOLOV5: Improved YOLOv5 Base on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. 2021 *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, 11-17 October 2021, 2778-2788.
<https://doi.org/10.1109/ICCVW54120.2021.00312>
- [10] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New States-of-the-Art for Real-Time Object Detectors. <https://arxiv.org/abs/2207.02696>
- [11] Girshick R. (2015) Fast R-CNN. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 07-13 December 2015, 5-7. <https://doi.org/10.1109/ICCV.2015.169>
- [12] Ren, S., He, K., Girshick, R. and Sun, J. (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149.
<https://doi.org/10.1109/TPAMI.2016.2577031>