



文章编号:1671-251X(2018)11-0032-05

DOI:10.13272/j.issn.1671-251x.2018050068

基于深度卷积神经网络的井下人员目标检测

唐士宇¹, 朱艾春², 张赛¹, 曹青峰¹, 崔冉¹, 华钢¹

(1. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221008;

2. 南京工业大学 计算机科学与技术学院, 江苏 南京 211816)

摘要:针对以人为中心的井下视频监控模式存在持续时间受限、多场景同时监视困难、人工监视结果处理不及时等问题,提出了基于深度卷积神经网络的井下人员目标检测方法。首先将输入图片缩放为固定尺寸,通过深度卷积神经网络操作后形成特征图;然后,通过区域建议网络在特征图上形成建议区域,并将建议区域池化为统一大小,送入全连接层进行运算;最后,根据概率分数高低选择最好的建议区域,自动生成需要的目标检测框。测试结果表明,该方法可以成功检测出矿井工作人员的头部目标,准确率达到87.6%。

关键词:煤矿安全;井下人员目标检测;头部检测;深度学习;卷积神经网络;Faster R-CNN

中图分类号:TD67 文献标志码:A

网络出版地址:<http://kns.cnki.net/kcms/detail/32.1627.TP.20181026.1443.002.html>

Target detection of underground personnel based on deep convolutional neural network

TANG Shiyu¹, ZHU Aichun², ZHANG Sai¹, CAO Qingfeng¹, CUI Ran¹, HUA Gang¹

(1. School of Information and Control Engineering, China University of Mining and Technology,

Xuzhou 221008, China; 2. College of Computer Science and Technology, Nanjing University of

Technology, Nanjing 211816, China)

Abstract: In view of problems that human-centered video monitoring mode had limited duration, multiple scenes were difficult to monitor at the same time, and results of manual monitoring were not processed in time, target detection method of underground personnel based on deep convolutional neural network was proposed. Firstly, input image was scaled to a fixed size, and a feature map was formed after operation of deep convolutional neural network; then, a suggestion area was formed on the feature map through area suggestion network, the suggestion area was pooled into a unified size which was sent to full connection layer for operation; finally, the best suggestion area was selected according to probability score, and the required target detection box was automatically generated. The test results show that the method can successfully detect head of underground personnel with an accuracy rate of 87.6%.

Key words: coal mine safety; target detection of underground personnel; head detection; deep learning; convolutional neural network; Faster R-CNN

收稿日期:2018-05-22;修回日期:2018-09-23;责任编辑:胡娴。

基金项目:国家自然科学基金项目(51574232)。

作者简介:唐士宇(1995—),男,江苏徐州人,硕士研究生,主要研究方向为深度学习、目标检测,E-mail:413255861@qq.com。通信作者:华钢(1963—),男,江苏徐州人,教授,博士,研究方向为矿山安全监控与监管、图像处理与模式识别,E-mail:ghua3323@163.com。

引用格式:唐士宇,朱艾春,张赛,等.基于深度卷积神经网络的井下人员目标检测[J].工矿自动化,2018,44(11):32-36.

TANG Shiyu, ZHU Aichu, ZHANG Sai, et al. Target detection of underground personnel based on deep convolutional neural network[J]. Industry and Mine Automation, 2018, 44(11): 32-36.

0 引言

我国煤矿安全事故一方面是由煤矿复杂地质与开采条件造成的,另一方面是由工作人员的不安全行为触发的^[1-2]。针对第2种情况,可以通过视频监控系统对煤矿不安全行为加以识别制止^[3-4],从而避免事故的发生。然而,目前以人为中心的井下视频监控模式存在持续时间受限、多场景同时监视困难、人工监视结果处理不及时等问题。因此,研究利用人工智能技术取代人工监视,实现井下人员不安全行为的主动识别,对保障煤矿安全具有重要意义。

要实现人员不安全行为的主动识别,首先要实现人员目标检测。传统的人员目标检测方法包括HOG特征法^[5]、Haar-Like特征法^[6]、LBP特征法^[7],这些方法特征层次很浅,现已被深度学习算法^[8-10]取代。在深度学习领域,基于深度卷积神经网络的算法^[11-12]具有特征层次深、检测准确率高、鲁棒性好的优点,已被逐步应用到各行业的检测任务中,但在煤矿领域的应用较少。本文提出了基于深度卷积神经网络的井下人员目标检测方法。由于在井下环境中身体部分经常被遮挡,而头部特征比全身特征更明显,故将全身目标检测转换为头部目标检测,以降低运算量,提高检测准确率。

1 深度卷积神经网络

自从基于区域的深度卷积神经网络(Region-based Convolutional Neural Networks, R-CNNs)在目标检测领域取得巨大成功以后,在其基础之上又衍生了快速的基于区域的深度卷积神经网络(Fast R-CNN)和更快速的基于区域的深度卷积神经网络(Faster R-CNN),其中Faster R-CNN具有更快的检测速度和更高的准确率,是当前研究的热点。Faster R-CNN^[13-15]总体网络结构如图1所示。首先将一张任意尺寸($P \times Q, P, Q > 0$)的输入图片缩放为固定尺寸($M \times N, M, N > 0$)的图片,缩放后的图片经过深度卷积神经网络操作后形成特征图;然后,通过区域建议网络在特征图上形成建议区域,并将建议区域池化为统一大小,送入全连接层进行运算;最后,根据概率分数高低选择最好的建议区域,自动生成需要的目标检测框。

1.1 VGG16网络模型

VGG16网络模型是一种较为流行的16层深度卷积神经网络模型,本文使用的VGG16网络模型包括13个卷积层、13个ReLU层和4个池化层。

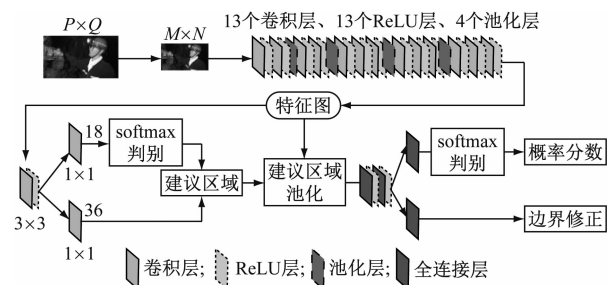


图1 Faster R-CNN 总体网络结构

Fig. 1 Overall network structure of Faster R-CNN

(1) 卷积层:对于一个多维的输入来说,每一维都有一个固定大小的卷积核,对该维进行卷积操作,形成多维输出的其中一维。VGG16网络模型中的13个卷积层都具有以下特征:为保证经过卷积运算后的图片尺寸不变,对输入矩阵的外围用数值0扩充,扩充后原图尺寸变为 $(M+2) \times (N+2)$,再用3像素 $\times$ 3像素大小的卷积核做卷积运算。

(2) ReLU层:激活函数使用以0为阈值的最大值激活函数,满足条件的正数传递给下一层,负数则直接归零。ReLU层紧跟在每一个卷积层后面。

(3) 池化层:使用最大值池化操作把小邻域内特征点的最大值整合成新的特征,有效减少了特征图的参数,起到了降维作用,降低了计算复杂度。

1.2 区域建议网络

原始输入图片经过VGG16深度卷积神经网络后,得到特征图,通过锚建立特征图到原始输入图片的反向联系,形成建议区域。锚的3种窗口面积尺寸为128像素 $\times$ 128像素,256像素 $\times$ 256像素,512像素 $\times$ 512像素。在每一种面积尺寸下,存在3种不同长宽比,分别为1:1,1:2,2:1,由此产生9种面积尺寸不同的锚。对深度卷积神经网络产生的特征图再进行一次卷积核大小为 3×3 的卷积,卷积窗口的中心点与原始图片中心点一一对应,根据9种大小各异的锚可以逆向推出原始图片的区域,也就得到了想要的建议区域。

为了让建议区域边界框更接近正确边界框,需要不断调优训练。锚得到的建议区域边界框和正确边界框对比如图2所示,其中大框是井下人员的标准边界框,小框为锚生成的边界框。Faster R-CNN模型认为小框为井下人员的头部,并没有完整地检测出井下人员的头部。

边界框回归如图3所示,框A代表锚生成的边界框,框G表示标准边界框。为了让框A更接近框G,采用平移缩放的思想进行变换:



图 2 锚生成的边界框和正确边界框对比
Fig. 2 Comparison of bounding box obtained by anchor and correct bounding box

$$G'_x = A_w d_{xA} + A_x \quad (1)$$

$$G'_y = A_h d_{yA} + A_y \quad (2)$$

$$G'_w = A_w \exp(d_{wA}) \quad (3)$$

$$G'_h = A_h \exp(d_{hA}) \quad (4)$$

式中: G'_x, G'_y, G'_w, G'_h 为框 A 变换后的中心点横坐标、纵坐标、框宽和框高; A_x, A_y, A_w, A_h 为框 A 变换前的中心点横坐标、纵坐标、框宽和框高; $d_{xA}, d_{yA}, d_{wA}, d_{hA}$ 是需要学习的变量。

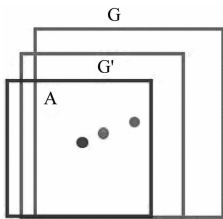


图 3 边界框回归
Fig. 3 Bounding box regression

如果框 A 和框 G 位置相差不大,可以认为对框 A 的变换是线性变换,表示式为

$$d_{*A} = \omega_*^T \cdot \Phi_A \quad (5)$$

式中: d_{*A} 为预测值; ω_* 为需要学习的参数; Φ_A 为特征图上锚对应的特征向量;下标“*”表示 x, y, w, h 。

1.3 建议区域池化网络

通过建议区域池化网络将建议区域的宽和高等分成 7 份,每个区域保留最大值,以达到将宽和高不同的建议区域标准化的目的,方便与分类网络中的全连接层相连接,形成特征向量。

1.4 分类网络

在特征图层面上,把已经池化的建议区域送入分类网络中的全连接层进行运算,得到特征向量。使用归一化指数函数 softmax 对特征向量进行计算,得出每个建议区域为井下人员头部的概率,即判别出该区域是否为井下人员。同时再次利用边界框回归获得每个建议区域的位置偏移量,进行边界修正,从而使目标检测框更加精确。

2 基于 Faster R-CNN 的煤矿井下人员头部检测

使用经 ImageNet 数据集预训练得到的初始化模型实现参数初始化。为了得到适用于煤矿井下环境的检测模型,需要制作煤矿井下人员头部数据集,并使用特定数据集进行优化训练。首先设定训练次数,达到训练次数后,判断损失函数是否收敛,若收敛则成功得到井下人员的头部检测模型,若不收敛则失败退出。整体流程如图 4 所示。

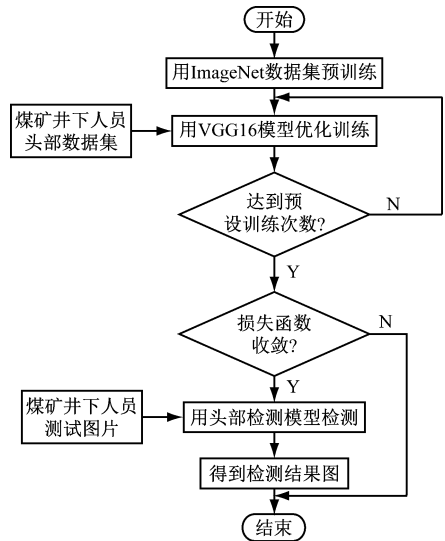


图 4 煤矿井下人员头部检测流程

Fig. 4 Detection flow of head of underground personnel

2.1 煤矿井下人员头部数据集的制作

由于井下环境特殊,用 ImageNet 数据集训练出的人员检测模型对于井下人员头部的检测效果不佳,所以,需要制作一个专用于煤矿井下人员头部检测的数据集用于模型调优,本文按照 VOC2007 数据集标准制作煤矿井下人员头部数据集。该数据集分为图片文件夹、标签文件夹、集合文件夹 3 个部分。图片文件夹下存放煤矿井下人员的图片,用于训练、验证和测试。标签文件夹下存放煤矿井下人员图片对应的标签文件,每个标签记录煤矿人员的头部坐标信息,标注示例如图 5 所示。在集合文件夹下包括 train.txt, trainval.txt, val.txt 和 test.txt 集合文件。其中 trainval.txt 集合文件中的图片数量占数据集总数的一半,剩下的图片放在 test.txt 集合文件中。将 trainval.txt 集合文件中的图片名分出一半的图片另存在 train.txt 集合文件中,剩下的图片另存在 val.txt 集合文件中。

2.2 VGG16 网络模型参数调整

根据需要识别的类别修改模型中的相关参数,



图5 LabelImg 软件标注示例

Fig. 5 Labeling example of LabelImg software
需要识别井下人员的头部这种类别标签和背景标签,所以输出的类别有2类。对如下文件进行修改。

在 py-faster-rcnn/models/pascal_voc/VGG16/faster_rcnn_alt_opt 文件夹中进行以下修改。

- (1) 修改 stage1_fast_rcnn_train.pt、stage2_fast_rcnn_train.pt 和 faster_rcnn_test.pt。
date 层: num_classes=2, 共2类数据。
cls_socre 层: num_output=2, 每种类别均有各自概率分数。

bbox_pred 层: (x,y,w,h) num_output=8, 每种类别均有4个位置信息。

- (2) 修改 stage1_rpn_train.pt 和 stage2_rpn_train.pt。

input-date 层: num_classes=2, 共2类数据。
在 py-faster-rcnn/lib/datasets 文件夹中, 修改 pascal_voc.py 文件, 添加 head 种类标签至 self._classes 中。

2.3 井下人员头部检测模型

检测模型的软硬件环境: 操作系统为 Ubuntu 系统, 深度学习框架为 caffe 框架, 处理器选用 NVIDIA GTX 1080Ti GPU, 学习率为 0.001, 迭代次数为 40 000。训练过程中损失函数值随训练次数变化曲线如图 6 所示。从图 6 可以看出, 当训练次数达到 28 000 时, 损失函数基本保持不变。井下人员头部检测模型达到收敛, 完成检测任务。

3 检测结果与分析

基于深度卷积神经网络的井下人员目标检测结果如图 7 所示, 图片来源于网络和视频截图, 方框是检测出来的矿井工作人员的头部区域。

使用相同的煤矿井下人员头部数据集、相同的 VGG16 模型, 采用 Faster R-CNN 算法和 Fast R-CNN 算法进行检测, 平均检测精度分别为 87.6% 和

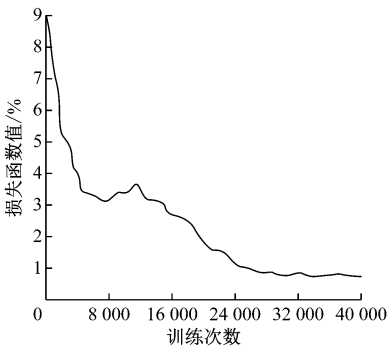


图6 损失函数值随训练次数变化曲线

Fig. 6 Curve of loss function value with changing of trocining numbers



图7 井下工作人员检测结果

Fig. 7 Test results of underground miners

61.0%, Faster R-CNN 算法准确率更高。除了本文使用的 VGG16 模型, 还有 ZF 模型, 考虑到模型之间的差异可能会对检测结果产生影响, 故使用相同的煤矿井下人员头部数据集、相同的 Faster R-CNN 算法、不同模型进行对比, 得到 VGG16 模型和 ZF 模型的平均检测精度分别为 87.6% 和 85.8%, 说明模型对检测结果的影响不大。在相同的 VGG16 模型、相同的 Faster R-CNN 算法下, 用井下人员头部数据集和 VOC2007 数据集进行对比测试, 平均检测精度分别为 87.6% 和 79.6%, 说明煤矿井下人员头部数据集比 VOC2007 数据集更适应于煤矿井下环境, 检测效果更好。

4 结论

- (1) 介绍了用深度卷积神经网络提取目标特征的原理及煤矿井下人员图像专用数据集制作流程, 并通过深度学习平台训练得到煤矿井下人员目标检

测模型。测试结果表明,用该模型可以成功检测出矿井工作人员的头部目标,准确率达到87.6%。

(2) 制约矿井工作人员头部目标检测精度的主要因素包括矿井中光线较暗、口罩遮挡、面部整洁度较差、拍摄角度变化大、图片分辨率低等。本文算法能够有效克服这些干扰因素,在低分辨率图片、黑白图片中也可以检测出井下人员的头部。但在单幅图片中有多名工作人员时,仍存在个别工作人员检测不出来和检测框不能完全包括头部的问题,接下来将对此进行进一步研究。

参考文献(References):

- [1] 王龙康,李祥春,李安金,等.我国煤矿安全生产现状分析及改善措施[J].中国煤炭,2016,42(9):96-100.
WANG Longkang, LI Xiangchun, LI Anjin, et al. Analysis and improvement measures on current situation of coal mine safety production in China[J]. China Coal, 2016, 42(9): 96-100.
- [2] 金永飞,靳运章.我国煤矿安全生产存在的若干问题及应对措施[J].煤矿安全,2015,46(4):234-236.
JIN Yongfei, JIN Yunzhang. Problems of coal mine safety production in China and theirs countermeasures [J]. Safety in Coal Mines, 2015, 46(4): 234-236.
- [3] 田将杉.煤矿视频监控系统的现状与发展趋势[J].机械管理开发,2017,32(9):130-131.
TIAN Jiangshan. Status and development trend of video monitoring system in coal mine machinery[J]. Mechanical Management and Development, 2017, 32(9): 130-131.
- [4] 张谢华.煤矿智能视频监控系统关键技术的研究[D].徐州:中国矿业大学,2013.
- [5] 田仙仙,鲍泓,徐成.一种改进HOG特征的行人检测算法[J].计算机科学,2014,41(9):320-324.
TIAN Xianxian, BAO Hong, XU Cheng. Improved HOG algorithm of pedestrian detection[J]. Computer Science, 2014, 41(9): 320-324.
- [6] 王庆伟,应自炉.一种基于Haar-Like T特征的人脸检测算法[J].模式识别与人工智能,2015,28(1):35-41.
WANG Qingwei, YING Zilu. A face detection algorithm based on Haar-Like T features[J]. Pattern Recognition and Artificial Intelligence, 2015, 28(1): 35-41.
- [7] 何云,吴怀宇,钟锐.基于多种LBP特征集成学习的人脸识别[J].计算机应用研究,2018,35(1):292-295.
HE Yun, WU Huaiyu, ZHONG Rui. Face recognition based on ensemble learning with multiple LBP features [J]. Application Research of Computers, 2018, 35(1): 292-295.
- [8] 尹宝才,王文通,王立春.深度学习研究综述[J].北京工业大学学报,2015,41(1):48-59.
YIN Baocai, WANG Wentong, WANG Lichun. Review of deep learning [J]. Journal of Beijing University of Technology, 2015, 41(1): 48-59.
- [9] 孙志远,鲁成祥,史忠植,等.深度学习研究与进展[J].计算机科学,2016,43(2):1-8.
SUN Zhiyuan, LU Chengxiang, SHI Zhongzhi, et al. Research and advances on deep learning[J]. Computer Science, 2016, 43(2): 1-8.
- [10] 郑胤,陈权屹,章毓晋.深度学习及其在目标和行为识别中的新进展[J].中国图象图形学报,2014,19(2):175-184.
ZHENG Yin, CHEN Quanqi, ZHANG Yujin. Deep learning and its new progress in object and behavior recognition[J]. Journal of Image and Graphics, 2014, 19(2): 175-184.
- [11] 段萌.基于卷积神经网络的图像识别方法研究[D].郑州:郑州大学,2017.
- [12] 卢宏涛,张秦川.深度卷积神经网络在计算机视觉中的应用研究综述[J].数据采集与处理,2016,31(1):1-17.
LU Hongtao, ZHANG Qinchuan. Applications of deep convolutional neural network in computer vision [J]. Journal of Data Acquisition and Processing, 2016, 31(1): 1-17.
- [13] 曹诗雨,刘跃虎,李辛昭.基于Fast R-CNN的车辆目标检测[J].中国图象图形学报,2017,22(5):671-677.
CAO Shiyu, LIU Yuehu, LI Xinzhaoh. Vehicle detection method based on Fast R-CNN[J]. Journal of Image and Graphics, 2017, 22(5): 671-677.
- [14] 梅舒欢,闵巍庆,刘林虎,等.基于Faster R-CNN的食品图像检索和分类[J].南京信息工程大学学报(自然科学版),2017,9(6):635-641.
MEI Shuhuan, MIN Weiqing, LIU Linhu, et al. Faster R-CNN based food image retrieval and classification [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2017, 9(6): 635-641.
- [15] 胡炎,单子力,高峰.基于Faster-RCNN和多分辨率SAR的海上舰船目标检测[J].无线电工程,2018,48(2):96-100.
HU Yan, SHAN Zili, GAO Feng. Ship detection based on Faster-RCNN and multiresolution SAR[J]. Radio Engineering, 2018, 48(2): 96-100.