

基于深度学习方法的胸片异物检测

侯鹏飞^{1,4}, 沈秀明², 袁明远³, 孙九爱⁴

1. 上海理工大学健康科学与工程学院, 上海 200093; 2. 上海市松江区卫生人才培训中心, 上海 201600; 3. 上海健康医学院附属周浦医院放射科, 上海 201318; 4. 上海健康医学院医学影像学院, 上海 201318

【摘要】为了提高胸片异物自动检测的能力,采用深度学习网络高效提取各种尺度和形状的异物影像特征,实现胸片中多种异物的自动稳定检测。在网络构建过程中,根据异物的影像特征,改进YOLO v4目标检测网络,通过在特征提取网络CSPDarkNet53中加入SE-block(Squeeze and Excitation),使模型能够区别利用各个通道的信息。实验结果表明,改进的深度学习网络在异物检测中能够实现92%的精确率和83%的召回率。因此,新的深度学习方法可用于胸片异物检测等应用场景,客观评判摄影质量,为放射影像的质量控制打下基础。

【关键词】胸片;异物检测;深度学习;YOLO网络

【中图分类号】R318;TP18

【文献标志码】A

【文章编号】1005-202X(2021)12-1518-06

Detecting foreign objects in chest radiographs based on a deep learning method

HOU Pengfei^{1,4}, SHEN Xiuming², YUAN Mingyuan³, SUN Jiulai⁴

1. School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 2. Healthcare Professional Training Center of Songjiang District, Shanghai 201600, China; 3. Department of Radiology, Zhoupu Hospital Affiliated to Shanghai University of Medicine and Health Sciences, Shanghai 201318, China; 4. School of Medical Imaging, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China

Abstract: In order to improve the ability of automatically detecting foreign objects in chest radiographs, a deep learning network is used to capture the imaging features of foreign objects with various scales and shapes efficiently, thereby realizing the automatic and stable detection of various foreign objects in chest radiographs. During network construction, according to the imaging features of foreign objects, YOLO v4 network is improved by embedding SE-block (Squeeze and Excitation) into the feature extraction network CSPDarkNet53 to make the model capable of distinguishing and utilizing the information of each channel. Experimental results demonstrate that the proposed deep learning network achieves 92% accuracy and 83% recall rate for foreign object detection. Therefore, the proposed deep learning approach can be used to identify foreign objects in chest radiographs and assess image quality automatically and objectively, which lays a foundation for the quality control of radiographic images.

Keywords: chest radiograph; foreign object detection; deep learning; YOLO network

前言

近年来随着数字X线摄影的广泛使用,X线摄影检查的数量和影像质量得到了很大的提高,已成为放射影像诊断检查最常用的方法^[1],但临床数据显示,数字X线摄影仍存在5%~15%的废片率^[2-3]。废

片的产生不仅降低了医疗资源的使用效率和影像诊断质量,后续补偿重复摄影还会给被检者增加不必要的电离辐射。为了利用最低辐射剂量获得最佳品质的医学影像,充分满足临床诊断的需要,X线摄影的甲级片率监控正成为医院放射科质量控制的一个重要任务。

胸部正位摄影是临床采用最为广泛的常规影像检查项目,也是放射影像质量控制检查项目之一。影响X线胸片质量的主要原因有摄影体位不正、呼吸气屏气不合理和被检者穿戴各种异物等。现阶段放射影像的质量控制主要通过人工检查阅片的方式核对技师操作的规范性。这种人工检查方法存在主观

【收稿日期】2021-07-14

【基金项目】上海健康医学院协同创新重点专项(SPCI-18-17-001)

【作者简介】侯鹏飞,硕士,研究方向:医学图像处理技术,E-mail: 892794685@qq.com

【通信作者】孙九爱,博士,副教授,研究方向:医学成像与图像处理技术,E-mail: sunja@sumhs.edu.cn

性强和效率低下等问题,而采用计算机辅助图像分析的方法可以实现影像质量的自动筛查,有效解决上述问题。异物的存在在X线影像中具有可解释和辨别的影像特征,因此本文以异物检测为研究目标,探索放射影像中异物的自动客观检测方法,为后续 的影像质量控制提供技术基础。

X线胸片中的异物检测可以归结为图像处理中的目标检测问题。Xue等^[4]最先提出对胸片中的异物进行自动检测,他们先对胸片中肺部区域进行分割,然后分别使用Viola-Jones和霍夫圆检测技术搜索线形或圆形特征,实现对纽扣类目标的检测。Zohora等^[5]首先使用几种不同的边缘检测算子对图像进行滤波处理,然后利用形态学方法得到候选结果,最后利用霍夫圆检测算法进行检测确认。最近Santosh等^[6]提出归一化互相关算法对胸片中的圆形异物进行初步检测得到候选物体,然后用无监督聚类的方法最后判断是否存在圆形异物。上述胸片异物检测研究都利用纽扣异物的圆形或其他能够标识物体的基本影像特征,对某一种异物的检测准确度可以达到80%左右^[6]。但实际临床摄影工作中异物种类各异,且影像特征变化繁多,尺度不一,基于特定的影像特征进行检测的方法很难取得满意的检测效果。王继元等^[7]在DR胸片质控体系研究中用Inception-V4网络对胸片中的异物进行分类,分类效果基本符合要求。但该方法主要实现异物的类型判断,为了确定异物位于肺野的具体位置,尚需要配合肺野的分割结果才能进行精准定位。

近年来发展起来的基于图像深度学习的方法突破了传统图像处理算法的缺陷,通过多通道和多尺度的卷积运算,可以抽取图像中具有不同形状和尺度的复杂影像特征,实现感兴趣目标的自动检测。目前基于深度学习方法的目标检测算法可分为Two-stage(两步法)和One-stage(单步法)两大类。Two-stage是把目标检测分为两个阶段,第一个阶段获得图像中若干个可能的物体候选框,然后对候选框进行分类和回归运算,最后得到物体的种类和位置信息,这种方法主要有R-CNN^[8]、Fast R-CNN^[9]、Faster R-CNN^[10]。而One-stage方法不需要产生候选框,直接对目标区域进行分类和回归,得到物体的种类和位置信息。One-stage方法比Two-stage方法大大提高检测效率,同时也提高精度,是目前目标检测的主流方法,主要有SSD^[11]、YOLO^[12]、YOLO v2-v4^[13-15]等。

YOLO目标检测网络通过将图像分为不同大小的子区域进行检测,这大大提高了检测速度,其中YOLO v4采用的CSPDarkNet53特征提取网络由若干个残差块组成,可以融合不同层次的信息提取有

用特征,分别对52×52、26×26和13×13尺度的特征层分别进行独立预测,实现多尺度的物体检测^[15]。在胸片异物检测的实验中可以发现,该网络结构对于形状模糊的异物检测精确率不高,因此本文在CSPDarkNet53特征提取网络中加入SE-block^[16],以区别对待不同通道信息的贡献程度,从而抑制贡献程度较低的信息,并增强贡献程度较高的信息。实验结果表明这种改进可以有效地提高模型的检测能力。

1 基于YOLO v4的异物检测

YOLO v4通过将YOLO v3的骨干特征提取网络DarkNet53改成CSPDarkNet53,并使用SPP^[17]结构和4个不同尺度卷积核进行最大池化处理,极大地增加了感受野,同时还引入PANet^[18]使网络能够反复利用图像特征,因此YOLO v4在检测精度和检测速度上都有了很大提升^[15]。

1.1 YOLO v4网络

YOLO v4主要由特征提取网络CSPDarkNet53、特征融合模块SPP和PANet、预测网络构成。其中CSPDarkNet53由5个残差块构成,残差块可以防止由于网络过深造成的过拟合等问题,最大化地提取图像的有用信息。SPP结构由4个卷积核大小为1×1、5×5、9×9和13×13的最大池化层构成,增加了感受野,能够更好地分离出上下文特征。PANet输入为CSPDarkNet53和SPP不同层的输出,通过对输入进行特征的反复提取,最后得到3个输出结果输入到最后的预测网络。YOLO在预测过程中先初始化9个不同大小的Anchor Box,在模型预测过程中相对于Anchor Box产生4个偏移参数 t_x 、 t_y 、 t_w 、 t_h ,如图1所示。

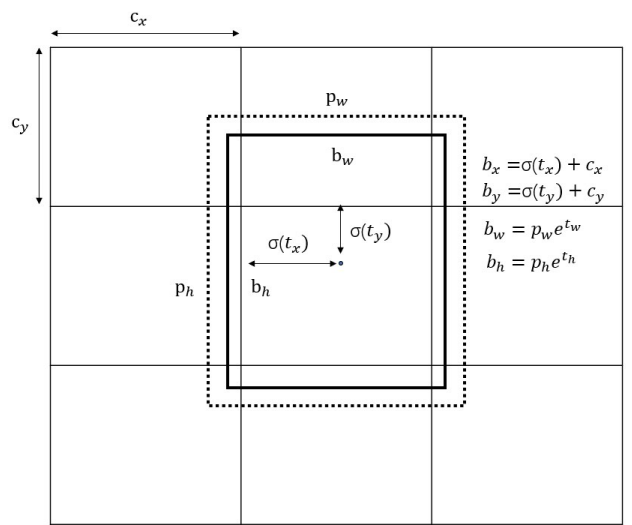


图1 Anchor Box和BoundingBox关系
Fig.1 Relationship between Anchor Box and BoundingBox

BoundingBox 调整公式为:

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

其中, b_x 和 b_y 为 BoundingBox 中心点横坐标和纵坐标, b_w 和 b_h 为 BoundingBox 的宽度和高度, c_x 和 c_y 是中心点坐标对应左上角方框的坐标, p_w 和 p_h 为 Anchor Box 对应的宽度和高度, σ 是 sigmoid 激活函数。

在计算 t_x, t_y, t_w, t_h 过程中, YOLO v4 采用 CIOU^[19] 作为回归优化的损失函数。CIOU 把 Ground Truth Box 与 BoundingBox 之间的重合率, 距离和尺度这些因素都考虑在内, 使得预测结果更加稳定和准确。CIOU 公式如下所示:

$$\text{CIOU} = \text{IOU} - \frac{\rho^2(b, b^{\text{gt}})}{C^2} - \alpha v \quad (5)$$

$$\text{IOU} = \frac{B \cap B^{\text{gt}}}{B \cup B^{\text{gt}}} \quad (6)$$

式(6)中, B 代表 BoundingBox, B^{gt} 代表 Ground Truth Box。式(5)中 b 代表 BoundingBox 中心点坐标, b^{gt} 代表 Ground Truth Box 中心点坐标。 $\rho^2(b, b^{\text{gt}})$ 代表 BoundingBox 和 Ground Truth Box 中心点之间的欧氏距离。 c 表示能够包含 BoundingBox 和 Ground Truth

Box 的最小 Box 的对角线长度。式中 α 和 v 可通过以下公式计算:

$$\alpha = \frac{v}{1 - \text{IOU} + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

式(8)中, w 和 h 代表 BoundingBox 的宽度和高度, w^{gt} 和 h^{gt} 代表 Ground Truth Box 的宽度和高度。最后得到损失函数为:

$$\text{loss}_{\text{CIOU}} = 1 - \text{IOU} + \frac{\rho^2(b, b^{\text{gt}})}{C^2} + \alpha v \quad (9)$$

1.2 FOXNet 网络结构

卷积神经参数和深度过多会很大程度增加检测物体的时间, 为了不给网络添加过多参数并且提高胸片异物检测的效果, 采用注意力机制对网络进行优化。由于 YOLO v4 骨干特征提取网络 CSPDarkNet53 在特征提取过程中不同的通道权重设置都相同, 不同通道的有用信息没有合理地利用起来。在异物检测神经网络的构建中, 我们改进了 YOLO v4, 在 CSPDarkNet53 网络的每个残差块中都加入 SE-block, 为每个通道使用不同的权重, 这样可以区别利用各个通道的信息^[16]。图2所示是对 CSPDarkNet53 中残差块加入 SE-block 后的网络, 称之为异物检测网络 FOXNet。

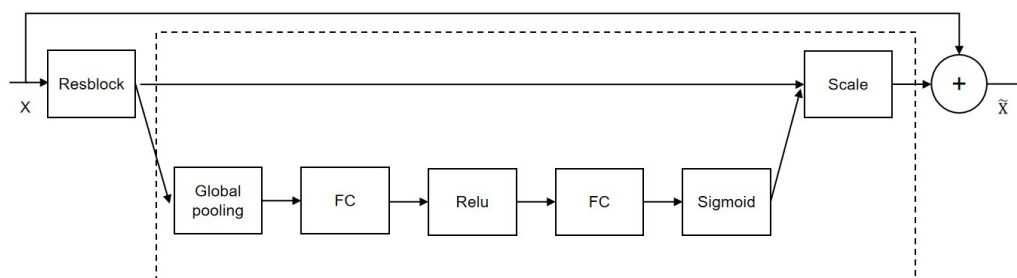


图2 改进后的残差模块

Fig.2 Improved residual network

图2中虚线框给出了残差网络中加入的 SE-block。网络经过残差块后有两条路径, 路径一直接连接到 Scale, 路径二先经过全局池化对网络进行压缩, 使得原本每个通道上的二维特征变为一个实数, 这个实数具有二维特征的全局感受野, 用来代表这个通道的特征, 经过压缩后整个网络尺度变为 $c \times 1 \times 1$; 然后网络经过两层全连接层建模通道间的相关性, 为每个特征通道生成权重, 权值经过学习后用来代表每个通道的重要程度, 得到的权重经过 Sigmoid 函数归一化为 0~1 之间的值。最后路径二产生的每个通道的归一化后的权重经过 Scale, 把权重加权在路

径一输出的每个通道上。

2 实验及结果分析

2.1 实验环境

开展本文实验所使用的电脑主要配置为 ubuntu18.04 操作系统, CPU 为 Inter(R) Xeon(R), 显卡为 TITAN X, 内存 64G。实验是通过 Python 语言编程实现, 应用深度学习框架 Pytorch 1.7.1 实现卷积神经网络的搭建。

2.2 数据及预处理

数据集取自 JF 医疗和 NVidia 公司举办的异物

检测比赛提供的共计 10 000 幅胸片图像,其中 5 000 幅为带有异物的胸片图像,5 000 幅为不带有异物的正常合格图像。图像中的肺野和异物由具有 1~3 年工作经验的共 12 位放射医生进行标注。实验中数据集被分为训练集、验证集和测试集,其中训练集包含 8 000 幅图片,验证集和测试集各包含 1 000 幅图片。图 3 所示为数据集的样本,图 3a 是不存在异物的正常图片,图 3b~图 3d 是包含异物的图片。本文进一步使用实际临床数据集验证算法的有效性和鲁棒性,从合作医院获取实际临床数据共 500 幅胸片,其中包含异物的为 50 幅,正常和异常的胸片数量比例为 9:1。

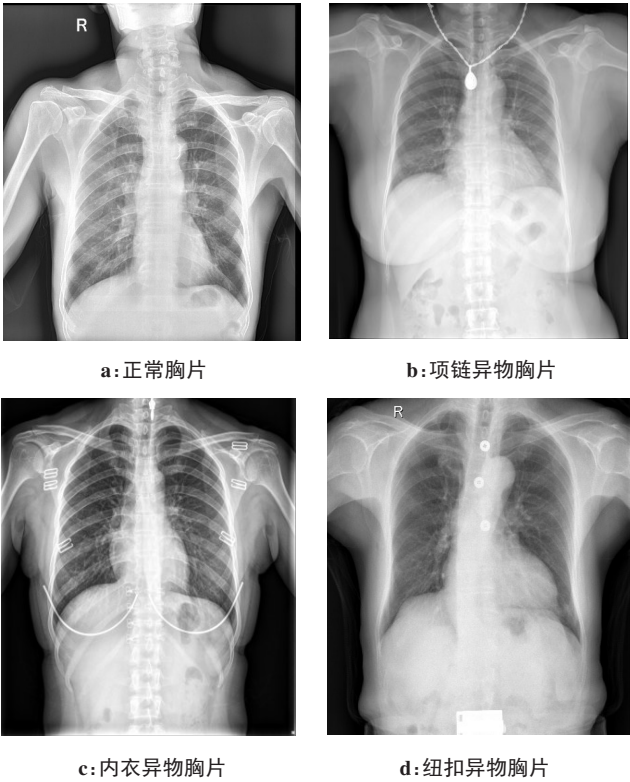


图3 正常数据和含有异物数据样本
Fig.3 Radiographs with and without foreign objects

在训练网络之前需要对数据进行预处理,这包括对图像的边 缘进行填充使图像变成正方形,并缩放成 416×416 像素大小;对数据集进行翻转、平移、拉伸变形和加入随机噪声,扩增数据集的大小,提高训练网络模型的鲁棒性。

2.3 评价标准

为了量化表达网络模型对于胸片异物检测的精度,采用准确率(Accuracy)、召回率(Recall)、精确率(Precision)、F1 score 和受试者工作特征曲线(ROC)^[20]等指标作为评价标准。其定义如下所示:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN}$$

(10)

$$Recall=\frac{TP}{TP+FN}$$

(11)

$$Precision=\frac{TP}{TP+FP}$$

(12)

$$F1\ score=2\frac{Recall\times Precision}{Recall+Precision}$$

(13)

其中,TN代表胸片图像中不存在异物,预测结果图像中也显示不存在异物;FP代表胸片图像中不存在异物,预测结果图像中显示存在异物;FN代表胸片图像中存在异物,预测结果图像中不存在异物;TP代表胸片图像中存在异物,预测结果图像中存在异物。

2.4 实验方法和结果

利用同样的数据集,实验分别对 Faster R-CNN、YOLO v3、YOLO v4 和 FOXNet 网络进行训练和测试,相应的精确率、召回率、准确率和 F1 score 结果如表 1 所示。

表 1 实验结果对比

Tab.1 Comparison of experimental results

网络模型	精确率	准确率	召回率	F1 score
Faster R-CNN	0.76	0.82	0.82	0.79
YOLO v3	0.88	0.84	0.77	0.82
YOLO v4	0.87	0.85	0.83	0.85
FOXNet	0.92	0.88	0.83	0.87

从表 1 的对比实验结果可以看出,利用典型两步法的 Faster R-CNN 获得的精确率和准确率都是最低的,分别为 76% 和 82%,但召回率较好,这说明两步法获得的敏感度较好,而特异性一般。相反,单步法的精确率和准确率都较好,都高于两步法,且 YOLO v4 和 FOXNet 的召回率也有所改善,达到 83%。YOLO v4 和 FOXNet 在 YOLO v3 基础上使用 SPP 结构,增加了感受野,还引入了 PANet 反复利用图像特征,使得真阳性预测的数量从 386 增长到 417,提高正样本的检出率,召回率显著提高 6%。FOXNet 在 YOLO v4 特征提取网络的每个残差块中加入 SE-block,使每个通道产生对应权重,充分利用各个通道感兴趣的信息,使得假阳性的数量从 61 降低到 34,有效降低误检率,精确率提高 5%。FOXNet 的综合性能指标 F1 score 是最高的,这也可以从各个网络在测试集测试获得的 ROC 曲线结果发现。如图 4 所示,Faster R-CNN、YOLO v3、YOLO v4 和 FOXNet 这 4 种网络获得的 ROC 曲线的 AUC 值分别为 0.846、0.904、0.922 和 0.945。

为了验证算法的有效性,本文利用合作医院提供的实际临床数据进行测试。测试结果显示,

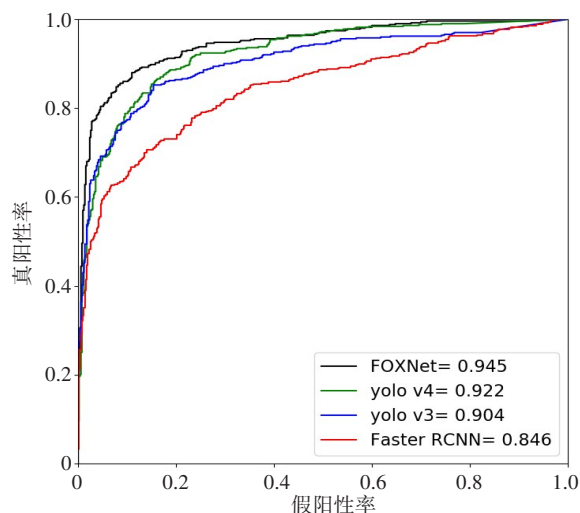


图4 不同网络结构ROC曲线示意图
Fig.4 ROC curves of different networks

FOXNet在实际临床测试数据中能正确检测出包含异物的42幅胸片,漏检8幅胸片,召回率达到84%。该结果比在异物检测比赛测试数据集上获得的结果还要好,这说明FOXNet能够可靠地检测出含有异物的胸片,具有一定的实际应用价值。图5所示为部分临床数据的检测结果。

图5a和图5b中的异物为植入性医疗电子或金属设备。虽然用来训练的数据集中异物主要为项链、内衣和纽扣等,但FOXNet也能很好地检测出训练过程中没有出现过的异物。图5c和图5d为图像像素仅为256×210低分辨率的情况,FOXNet网络能以0.83和0.80的置信度检测出不是很明显的异物。图5e~图5h中的异物(衣服、装饰等)是肉眼难以观察到的,FOXNet网络依然能够准确识别。这些结果表明,

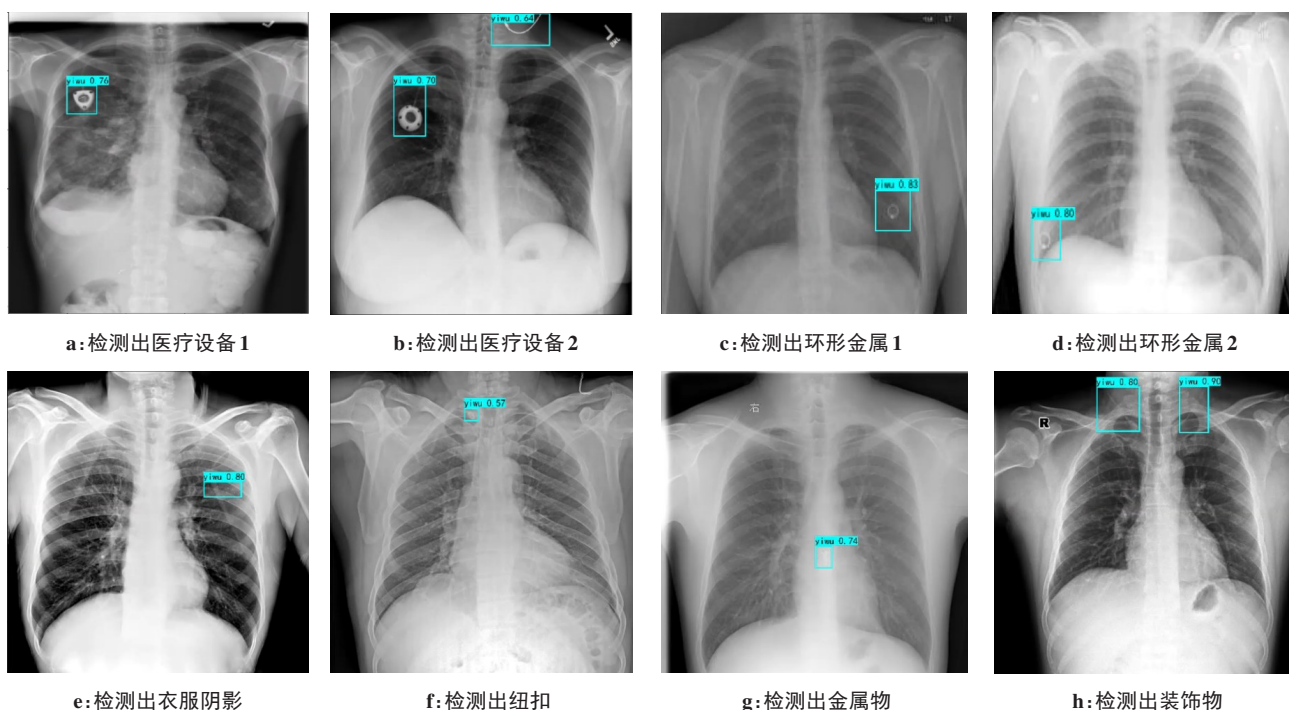


图5 临床胸片检测样例
Fig.5 Examples of foreign objects detected from clinical chest radiographs

FOXNet有着很好的异物检测泛化能力,能够用于胸片的异物自动检测的实际应用场景。

3 讨论

以往对于胸片中的异物检测的传统图像处理方法,局限于检测纽扣或者具有特定形状的某一类物体,难以适用于临床胸片中存在的尺寸和形状变化繁多的异物检测任务。本研究提出用深度学习的方法对X线胸片中的异物进行检测,卷积神经网络中包含有多种不同尺寸和参数的卷积核,能够学习提取

图像在不同尺度下的特征,实验证明该方法能够克服传统图像处理的固有缺点。

本研究基于YOLO v4目标检测网络,并在网络中增加运用了注意力机制的SE-block,使得新的网络能够充分利用不同通道的信息。实验表明这样的改进能够有效提高异物检测的精确率和准确率,降低异物检测的误检率,使自动检测的结果更贴近摄影质量控制的真实需求,这对在放射摄影质量评估中推广使用基于深度学习的异物自动检测方法具有重要意义。

虽然FOXNet的实现改善了胸片异物检测的精确率和准确率,并在网络的总体检测性能上有所提高,但该方法的异物检测召回率依然与两步法相近。导致这个结果可能有多种原因,如用于训练和测试的图片是通过医学影像压缩获取的,其空间和灰度分辨率大大缩小,部分异物在压缩后很难辨别或甚至消失。因此,在异物检测的漏检率方面FOXNet还需要开展很多的研究,这也将是 我们下一步的研究目标。

4 结 论

为了实现放射摄影质量的客观评价,提高胸部正位片的优良率,避免被检者接受不必要的电离辐射,本文提出一种基于深度学习的方法用来自动检测和定位胸片中的异物。实验结果表明,该方法的异物检测可以实现92%的精确率和83%的召回率。

【参考文献】

[1] 施小珍. 胸部X线联合CT检查对于肺部感染早期筛查的诊断意义分析[J]. 实用心脑血管病杂志, 2019, 27(S1): 214-216.
SHI X Z. Diagnostic significance of chest X-ray combined with CT in early screening of pulmonary infection[J]. Practical Journal of Cardiac Cerebral Pneumal and Vascular Disease, 2019, 27(S1): 214-216.

[2] SAJJAD R, JALAL B, EHSAN S, et al. Reject analysis in digital radiography: a local study on radiographers and students' attitude in Iran[J]. Med J Islam Repub Iran, 2019, 33(1): 49.

[3] BJØRN H, BLOMBERG R T, CAMILLA J, et al. Image rejects in general direct digital radiography[J]. Acta Radiol Open, 2015, 4(10): 1.

[4] XUE Z, CANDEMIR S, ANTANI S, et al. Foreign object detection in chest X-rays[C]// IEEE International Conference on Bioinformatics and Biomedicine. IEEE, 2015: 956-961.

[5] ZOHORA F T, SANTOSH K C C. Foreign circular element detection in chest X-Rays for effective automated pulmonary abnormality

screening[J]. Int J Comput Vis, 2017, 7(2): 36-49.

[6] SANTOSH K C, ZOHORA F T. Circle-like foreign element detection in chest X-rays using normalized cross-correlation and unsupervised clustering [C]//SPIE Medical Imaging Conference (Machine Learning). 2017.

[7] 王继元, 李真林, 蒲立新, 等. 基于人工智能的正位DR胸片质控体系研究与应用[J]. 生物医学工程学杂志, 2020, 37(1): 158-168.
WANG J Y, LI Z L, PU L X, et al. Research and application of orthotopic DR chest radiograph quality control system based on artificia intelligence[J]. Journal of Biomedical Engineering, 2020, 37 (1): 158-168.

[8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Conference on Computer Vision and Pattern Recognition. 2014: 580-587.

[9] GIRSHICK R. Fast R-CNN[C]//IEEE International Conference on Computer Vision. 2015: 1440-1448.

[10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Trans Pattern Anal, 2017, 39(6): 1137-1149.

[11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[J]. arXiv Preprint, 2015, arXiv: 1512.02325.

[12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Computer Vision & Pattern Recognition. IEEE, 2016.

[13] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// Computer Vision & Pattern Recognition. IEEE, 2017: 6517-6525.

[14] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv Preprint, 2018, arXiv: 1804.02767.

[15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv Preprint, 2020, arXiv: 2004. 10934.

[16] JIE H, LI S, SAMUEL A, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(8): 2011-2023.

[17] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-level performance on imagenet classification [C]// IEEE International Conference on Computer Vision. 2015: 1026-1034.

[18] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[J]. arXiv preprint, 2018, arXiv: 1803.01534.

[19] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. arXiv Preprint, 2020, arXiv: 1911.08287.

[20] LUSTED, LEE B. Logical analysis in Roentgen diagnosis [J]. Radiology, 1960, 74(2): 178.

(编辑:陈丽霞)