

基于改进 SSD 的 X 光图像管制刀具检测与识别

郭瑞鸿*, 张莉, 杨莹, 曹洋, 孟俊熙

西安工程大学电子信息学院, 陕西 西安 710048

摘要 以安检 X 光图像管制刀具自动检测识别系统为研究对象, 针对原始 SSD(Single Shot MultiBox Detector)算法对浅层特征图表征能力不强, 在训练阶段小目标特征逐渐消失, 检测精度与实时性不佳, 存在对安检危险品中管制刀具等小目标漏检误检等问题, 从两个方面对原始 SSD 进行改进: 一方面, 用抗退化性能更强的 ResNet34 网络替换 SSD 中的基础网络 VGG16, 构建 SSD-ResNet34 网络模型, 对基础网络后三层作卷积并进行轻量级网络融合, 形成新的低层特征图; 将网络部分扩展层作反卷积, 形成新的高层特征图。另一方面, 采用跳跃连接的方式将高层特征图和低层特征图进行多尺度特征融合。经实验分析, 改进后的算法对 X 光图像管制刀具等小目标的检测精度和速度均有明显提升, 且算法鲁棒性好, 实时性良好。在 VOC2007+2012 通用数据集上, 改进 SSD 算法的检测精度比 SSD 算法高 1.7%, 达到了 80.5%。

关键词 探测器; X 光图像; 深度学习; 目标检测; 特征融合; 残差神经网络

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0404001

X-Ray Image Controlled Knife Detection and Recognition Based on Improved SSD

Guo Ruihong*, Zhang Li, Yang Ying, Cao Yang, Meng Junxi

College of Electronics and Information, Xi'an Polytechnic University, Shaanxi, Xi'an 710048, China

Abstract In the automatic X-ray imaging systems used to identify knives in security check, using the original single shot multibox detector (SSD) algorithm, the shallow feature maps are poorly represented, features of small targets gradually disappear during the training stage, leading to low detection accuracy and poor real-time performance, and the small targets such as the controlled knives in security check are missing and checked out by mistake. To solve this problem, the original SSD was improved in two ways. On the one hand, the SSD-Resnet34 network model was constructed by replacing the basic network VGG16 in the SSD using a ResNet34 network with stronger anti-degradation performance, and the last three layers of the basic network were convolved and a new low-level feature map was created by lightweight network fusion. Part of the extended layer of the network was deconvolved to form a new high-level feature map. On the other hand, jumping connection was adopted to achieve multi-scale feature fusion between the high-level feature map and the low-level feature map. Analysis of test data shows that the improved algorithm demonstrates improved detection speed and detection accuracy of small targets, such as the X-ray image controlled knives. And the algorithm demonstrates improved robustness and high real-time performance. Using the VOC2007+2012 general dataset, the detection accuracy of the improved SSD algorithm is 1.7% higher than that of the SSD algorithm, reaching 80.5%.

Key words detectors; X-ray image; deep learning; target detection; feature fusion; residual neural network

OCIS codes 040.7480; 150.0155; 150.1135

收稿日期: 2020-09-10; 修回日期: 2020-10-18; 录用日期: 2020-11-05

基金项目: 国家自然科学基金(51607133)、陕西省教育厅研究项目(10JK510)

* E-mail: rhguoo@qq.com

1 引言

目前, X 射线安检技术的非接触式安检设备是应用最广泛的工具, 被大量应用于物流领域, 以及交通枢纽和一些人员较为密集的公共场所。精准的管制物品自动检测技术是安检领域一个主要的研究方向。现有的基于深度学习的目标检测算法存在对小目标误检、漏检等问题^[1], 整体检测效果不佳。近年来基于卷积神经网络的算法被广泛应用^[2], 目前在深度学习目标检测领域主要有基于区域建议的两阶段系列算法 (R-CNN^[3]、Fast RCNN^[4]、Faster RCNN^[5]、R-FCN^[6]、FPN^[7]) 和基于回归问题的单阶段系列算法 [YOLO (You Only Look Once)^[8]、SSD (Single Shot Multibox Detector)]。SSD 算法较其他算法有良好的性能, 检测精度和速度平衡, 鲁棒性较强, 在小目标检测领域应用较为成熟, 但也有不足之处。为提高 SSD 算法对安检图像管制刀具小目标的检测精度, 需对传统 SSD 算法进行改进。

本文针对 X 光图像管制刀具检测与识别主要完成了以下三个方面的工作:

1) 基于原始 SSD 算法, 采用 ResNet34 网络替换 SSD 的 VGG16 基础网络, 对 ResNet34 基础网络后三层作卷积并进行轻量级网络融合, 形成新的三层特征图, 较好地解决了网络退化问题和梯度消失问题, 改善了 SSD 在小目标检测上的缺陷;

2) 将 SSD-ResNet34 网络部分扩展层作反卷积, 形成新的高层网络, 设计新高层网络与新低层网络跳跃连接的融合模块, 有效提高了 SSD 模型对小目标的检测性能;

3) 采集各类管制刀具的 X 光图像照片, 并进行人工标注, 制作管制刀具的专用数据集 Specialized Data set for Controlled Knife (SDCK)。

2 理论基础

2.1 问题分析

在 X 光图像管制刀具检测识别应用中, 由于 X 光图像管制刀具数据集中刀具图像相对于正常的摄影图像分辨率低、水果刀等小目标在原图中的尺寸较小, 故原始图像经神经网络下采样处理会导致小目标特征图尺寸大小基本为个位数像素, 分类网络对小目标的分类效果较差。X 光图像管制刀具数据集中小目标图像经大小调整后宽高分布被急剧挤压, 宽高比分布极不均匀, 导致检测网络中不同宽高比的候选框均无法有效接近真实值, 小目标特征感

受野的映射图将可能大于小目标在原始图像中的尺寸, 检测效果不佳。在 X 光图像数据集中, 复杂的无关行李属于背景, 网络模型在检测小目标时会受到过多背景噪声的影响, 网络模型对小目标的检测精度较低。传统 SSD 算法在检测小目标时使用低层特征图, 低层特征图因感受野小以及语义信息低等导致对小目标特征提取不充分, 训练阶段经过层层池化导致小目标特征消失, 故传统 SSD 算法达不到小目标检测精度的要求。

2.2 构建数据集

本文所用的检测数据集采集自高铁 X 射线安检机, 将实际采集数据图片与摆拍数据图片相融合, 构成了一定量的安检管制刀具图片库。对采集到的数据集图像采用旋转、水平翻转等方法进行扩充, 共得到 6872 张照片。利用 labelme 数据集标注工具对所得图像进行人工分类和标注。对采集得到的 X 光图像管制刀具图片进行标注时, 将图片中所有管制刀具尽可能全部标记, 如果图像内的管制刀具边界不全或不清晰, 则放弃对该样本的标记。采用统一标准对各类管制刀具目标进行标记, 确保所标记的外接矩形框和样本的类别信息准确无误。构建的数据集格式与 PASCAL VOC2007 形式一致, 利用 XML 存储每一张图像中对应目标的位置和类别信息。

数据集包括 7 种类别的危险刀具物品: 菜刀 (kitchen knife)、水果刀 (fruit knife)、单刃长刀 (hacking knife)、匕首 (dagger)、剪刀 (scissor)、扳手 (spanner)、小刀片 (little knife)。为防止网络过拟合, 在实际采集过程中训练样本采取人为增加的方式, 图 1 为数据集示例照片。



图 1 数据集示例照片

Fig. 1 Sample photos of the dataset

2.3 深度残差网络

在训练深度学习网络模型时, 随着训练神经网络深度的加深, 模型准确率达到饱和, 随后网络深度

继续增加,模型准确率开始下降。深度学习在神经网络深度变化过程中遇到的问题主要有梯度消失和梯度爆炸。传统数据初始化和正则化虽然解决了梯度消失的问题,但网络性能产生退化现象。而残差网络不仅解决了退化和梯度消失问题,还提升了网络性能。残差网络结构模块,可以很大程度地加深网络结构,且最终的分类效果较好。残差网络结构模块如图 2 所示。

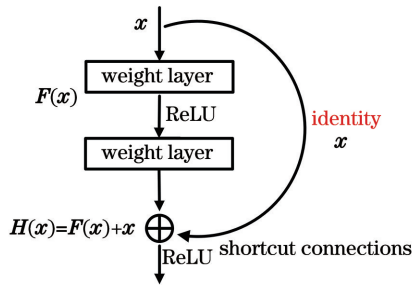


图 2 残差网络结构模块

Fig. 2 Residual network structure module

在图 2 的残差网络结构图中,通过捷径连接(shortcut connections)的方式,直接把输入 x 传到输出作为初始结果,输出结果为

$$F(x) = F(x) + x. \quad (1)$$

表 1 各个网络性能测试结果

Table 1 Results of each network performance test

Network	Speed / (frame · s ⁻¹)	Accuracy	Mode size / M	FLOPS / 10 ⁹
DenseNet121	114	74.3	32	1.9
VGG16	183	71.7	537	15.3
MobileNetv2	231	72.0	14	0.52
ShuffleNet	303	70.8	21	0.524
ResNet34	402	73.1	87	3.6

从表 1 可以看出,ResNet34 网络的速度和精度均比 SSD 原始基础网络 VGG16 高很多,并且从每秒浮点运算次数(FLOPS)来看,ResNet34 网络要比 VGG16 小很多,这表明 ResNet34 模型的时间复杂度远低于 VGG16。综上所述,本文选择了 ResNet34 作为改进 SSD 算法的基础分类网络。

对基础网络 ResNet34 的 Conv3_x、Conv4_x、Conv5_x 进行卷积处理,生成新的特征层,分别记为 Conv3_{x_0}、Conv4_{x_0}、Conv5_{x_0},新的特征层比原始特征层语义信息更丰富,更好地保留了原始特征图的边缘信息,且新的特征层维数不超过原始特征层维数^[15]。然后将 Conv3_x 和 Conv3_{x_0}、Conv4_x 和 Conv4_{x_0}、Conv5_x 和 Conv5_{x_0}

在网络学习达到较饱和准确率,即 $F(x)=0$ 时,学习目标转化为恒等映射的学习,残差网络输出变为 $F(x)=F(x)-x$ 。训练网络时残差结果接近于 0,使得较深网络训练能够确保准确率不降低。残差网络这种跳跃式结构模块,可以将某一层的输出跨越几层后直接作为后面某一层的输入,较好地解决了堆叠多层网络时梯度下降及网络性能退化的问题。

3 改进 SSD 算法及网络学习

针对原始 SSD 算法对 X 光图像管制刀具等小目标检测精度不高、网络存在性能退化等问题,对传统 SSD 算法进行改进:改进基础网络,优化侯框的回归和分类任务;对低层和高层不同特征图进行多尺度特征融合,提高模型对小目标检测的性能。

3.1 改进基础网络

在目标检测算法网络中分类网络的性能对其性能有重大影响,分类网络去掉全连接层和损失层就是基础网络部分。为验证常见基础分类网络的性能,基于 Caffe^[9] 框架在 ImageNet2012 数据集上对 DenseNet^[10]、VGGNet^[11]、MobileNetv2^[12]、ShuffleNet^[13] 和 ResNet^[14] 各个网络进行测试,结果如表 1 所示。

按图 3 的方式依次进行轻量级特征融合,得到 Conv3_{x_1}、Conv4_{x_1}、Conv5_{x_1} 三个特征层。

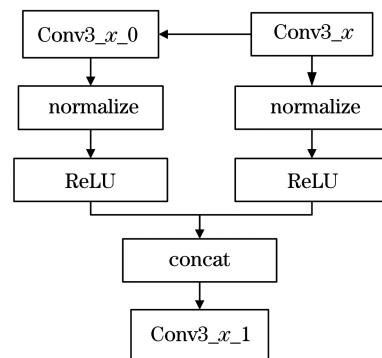


图 3 特征连接融合方式

Fig. 3 Feature connection and fusion mode

3.2 改进拓展层构建新的网络模型

借鉴 DSSD 网络和 MFDSSD 网络^[16]的思想,改进算法网络结构,如图 4 所示,虚线框内为 ResNet34 残差结构层,Conv5_2~Conv9_2 是新添加的拓展卷积层。通过改进基础网络后的低层网络和改进拓展层后高层网络之间的融合,发挥高层网络与低层网络的优势,使得模型对安检管制刀具等小目标检测性能得到提升。将 SSD-ResNet34 拓展层的 Conv6_2、

Conv7_2、Conv8_2 分别进行反卷积,得到改进网络的反卷积层分别为 Conv10_2、Conv11_2、Conv12_2。然后将反卷积后得到的高层网络 Conv10_2、Conv11_2、Conv12_2 分别和改进基础网络得到的低层网络 Conv5_x_1、Conv4_x_1、Conv3_x_1 进行跳跃连接的多尺度特征融合。低层网络和高层网络之间多尺度的特征融合充分利用特征图上下文语义信息,增强低层网络感受野,提高了高层网络几何信息的表征能力。

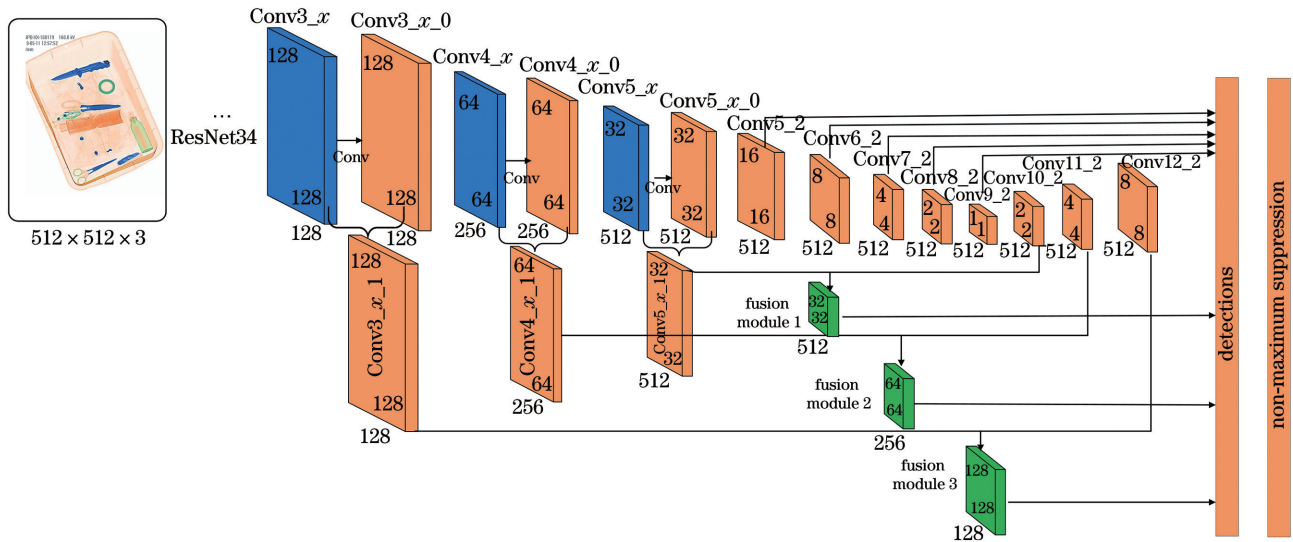


图 4 改进网络模型图

Fig. 4 Improved network model diagram

本文算法与 DSSD 和 MFDSSD 算法相比,优点在于本文改进算法模型的基础分类网络为 ResNet34, ResNet34 相比 DSSD 网络中基础分类网络 ResNet101,网络深度较浅,故改进后模型的复杂度降低,实时性更高。且本文改进算法对基础分类网络 ResNet34 后三层作相同卷积处理后,与原来特征层两两进行轻量级特征融合,得到新的 3 层特征图,使得基础网络对小目标提取的特征信息更加丰富,然后再与反卷积模块的高层特征图进行 3 个跳跃连接的特征融合。改进后网络模型

复杂度较低,通过低层和高层网络之间的跳跃式特征融合模式,使得模型进一步提升小目标的检测性能。由图 4 可以看出本文改进算法网络产生的多尺度特征图共有 8 个尺寸,SSD 和 MFDSSD 分别有 6 个和 7 个,所以本文改进算法同时兼顾了 DSSD 和 MFDSSD 的优势且特征图更具多样性,多样性的特征图对小目标的检测提供了丰富的上下文语义信息。本文改进算法有三个跳跃连接的特征融合模块,具体特征融合模块的融合策略如图 5 所示。

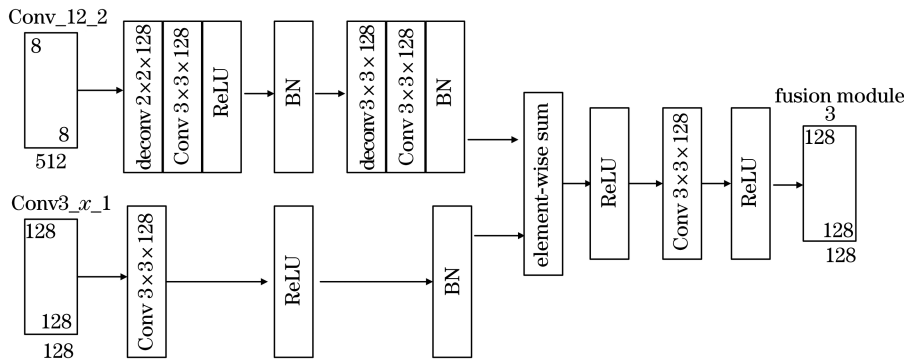


图 5 跳跃连接特征融合方式

Fig. 5 Jumping connection feature fusion pattern diagram

3.3 网络学习策略

在训练网络模型时建立真实标签和默认框的对应关系,从默认框中选择真实标签。首先将每个真实标签框与默认框中 Jaccard overlap^[17-18] 值最高的默认框相匹配,确保每个真实标签框有一个匹配的默认框,然后将默认框与真实标签框中 Jaccard overlap 重叠度高于阈值为 0.5 的默认框相匹配,这样就使得网络对多个重叠的默认框的预测置信度较高,而不是只选择重叠度最大的一个。Jaccard overlap 的表达式为

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (2)$$

式中: A, B 分别表示某种集合。

定位损失是预测框(l)和真实框(g)之间的 Smooth L1 损失,通过回归到默认框(d)相对于真实框中心(c_x, c_y)和宽(w)高(h)的偏移量来表达,表达式为

$$L_{Loc}(X, l, g) = \sum_{i \in P_{os}^m \in \{c_x, c_y, w, h\}} \sum X_{xi}^k \text{smooth}_{L1}(l_i^m - \hat{g}_i^m), \quad (3)$$

$$\hat{g}_j^{c_x} = (g_j^{c_x} - d_i^{c_x})/d_i^w,$$

$$\hat{g}_j^{c_y} = (g_j^{c_y} - d_i^{c_y})/d_i^h,$$

$$\hat{g}_j^w = \ln\left(\frac{g_j^w}{d_j^w}\right),$$

$$\hat{g}_j^h = \ln\left(\frac{g_j^h}{d_j^h}\right), \quad (4)$$

式中: P_{os} 为正样本; X_{xi}^k 为指示函数, k 为类别, j 为真实框序号; i 为预测框序号; N 表示匹配锚窗的数量,通过交叉验证设置 α 值为 1。

总的损失函数为分类损失(Conf loss, L_{conf})和定位损失(Loc loss)的加权平均,表达式为

$$L(f, c, l, g) = \frac{1}{N} [L_{Conf}(f, c) + \alpha L_{Loc}(f, l, g)], \quad (5)$$

式中: f 为指示函数; c 为置信度。

4 实验与分析

4.1 实验环境和参数配置

本文的实验过程都在 Linux PC 端上进行,训练模型和测试算法性能所使用的具体硬件和软件环境如表 2 所示。

训练模型的初始学习率(learning rate)设置为 0.001,迭代次数达到 40000 次后更改为 0.0001,继续迭代 40000 次以后模型训练结束。在

表 2 实验环境配置

Table 2 Experimental environment configuration

Hardware	Software
CPU: Intel i7-9750H CPU	Operating system
GTX 1660Ti-8G	Ubuntu16.04
GPU: NVIDIA RTX2080ti	Frame: Tensorflow
RAM: 11G×4	Language: Python

Tensorflow 架构上进行模型训练,每间隔 20000 次进行一次测试,优化器动量(momentum)和权重衰减率(weight decay)分别设置为 0.9000 和 0.0005,批大小(batch size)设置为 64,采用随机梯度下降法(SGD)优化参数。

4.2 方法对比分析

4.2.1 性能评估指标

目前通常使用均值平均精度(mAP)和检测速度(帧每秒, $\text{frame} \cdot \text{s}^{-1}$)来评价目标检测算法的性能。均值平均精度(mAP)由准确率 $P_{\text{precision}}$ 和召回率 R_{recall} 构成,表达式为

$$P_{\text{precision}} = T_p / (T_p + F_p), \quad (6)$$

$$R_{\text{recall}} = T_p / (T_p + F_N), \quad (7)$$

式中: T_p 表示真正类,即模型预测为正的样本; F_p 表示真负类,即模型预测为负的样本; F_N 表示假负类,即模型预测为正的负样本。

在目标检测算法中每个类别都可以得到准确率-召回率曲线(P-R 曲线),曲线与横坐标轴围成的面积就是平均精度(AP)值,求所有类别的 AP 的平均值,就是 mAP 值。为评价本文改进 SSD 算法的综合性能,在经典数据集 PASCAL VOC(2007+2012)和 X 光管制刀具自制数据集 SDCK 上对改进 SSD 进行训练和评估。其中 VOC 数据集中目标检测数据集含有 20 类目标图像,X 光管制刀具自制数据集 SDCK 数据集中含有 7 类刀具,均以 IOU(Intersection over Union)大于 0.5 为阈值评估得分。

4.2.2 检测精度对比

为验证本文改进的 SSD 算法性能,选择原始 SSD 算法、DSSD 算法、MFDSSD 算法和本文算法在 SDCK 管制刀具数据集上进行了 4 组测试,如表 3 所示。

表 3 各算法在 SDCK 数据集上的 mAP

Table 3 mAP of each algorithm on SDCK dataset

Method	SSD	DSSD	MFDSSD	Our method
mAP	89.8	91.3	90.5	92.6

unit: %

本文改进的 SSD 检测算法在 SDCK 管制刀具数据上相比经典的 SSD 算法精度提高 2.8 个百分点,且优于 MFSSD 和 DSSD 算法。本文算法利用残差网络 ResNet34 代替原 VGG16 作为 SSD 的基础网络,构建残差 SSD 检测器模型,在此基础上对 ResNet34 基础网络后三层作卷积并进行轻量级网络融合形成新的三层特征图,将 SSD-ResNet34 网

络部分扩展层作反卷积形成新的高层网络,设计新高层网络与新低层网络跳跃连接的融合模块,在 SDCK 管制刀具上的检测精度比 MFDSSD 算法提高 2.1 个百分点,可见本文改进的算法在一定程度上改善了对管制刀具小目标的检测性能。原始 SSD 算法在 SDCK 管制刀具数据集的上的 mAP 如表 4 所示。

表 4 SSD 算法在 SDCK 管制刀具数据集的上的 mAP

Table 4 mAP of SSD algorithm on SDCK controlled tool dataset

unit: %

Method	SSD	Kitchen knife	Fruit knife	Hacking knife	Dagger	Scissor	Spanner	Little knife
mAP	89.8	91.0	88.9	93.0	89.2	87.8	91.3	87.5

为验证本文改进算法的综合性能,接着在 VOC2007+2012 通用数据集上进行实验测试,测试结果如表 5。本文改进算法的优点如下。

表 5 各算法在 VOC2007+2012 数据集上的 mAP

Table 5 mAP of each algorithm on VOC2007+2012 dataset

Method	Train	Test	Input size	mAP / %
SSD	2007+2012	2007	300×300	78.8
DSSD	2007+2012	2007	321×321	80.3
MFDSSD	2007+2012	2007	300×300	80.0
Our method	2007+2012	2007	512×512	80.5

1) 本文改进算法的检测精度得到一定程度的提升,这说明本文改进方法的可行性较好;

2) 与传统 SSD 算法训练效果相比,本文改进算法的检测精度提高 1.7 个百分点,可见本文改进算法对 VOC2007+2012 通用数据集中普通物体的检测也有较好的性能体现;

3) 与 DSSD 算法相比,本文改进算法的检测精度提高 0.2 个百分点,本文改进算法在网络复杂度大大降低的基础上依然能够保持良好检测精度,检测速度优于 DSSD,更加满足实际应用中的实时性要求;

4) 与 MFDSSD 算法相比,本文改进算法在通用数据集上的检测精度比 MFDSSD 算法高 0.5 个百分点,本文改进算法利用基础网络对小目标提取的特征信息更加丰富,网络产生的多尺度特征图共有 8 个尺寸,MFDSSD 有 7 个,所以本文改进算法对小目标的检测精度稍优于 MFDSSD。

4.2.3 检测速度对比

为测试本文改进 SSD 算法的检测速度,选择 SSD 算法、DSSD 算法、MFDSSD 算法,搭载两台 NVIDIA RTX2080ti 的 GPU 服务器在 SDCK 管

制刀具数据集上进行测试,得到每个算法的检测速度(帧每秒,frame·s⁻¹),各个算法与本文算法的检测速度对比如表 6 所示。

表 6 各算法在 SDCK 数据集上的检测速度

Table 6 Detection speed of each algorithm on SDCK dataset

unit: frame·s⁻¹

Method	SSD	DSSD	MFDSSD	Our method
Speed	18.3	9.6	13.5	16.7

从表 6 中的对比数据可以看出,本文改进算法在 SDCK 管制刀具数据集上的检测速度为 16.7 frame·s⁻¹,与原始 DSSD 算法的检测速度相比提高 7.1 frame·s⁻¹,比 MFDSSD 提高 3.2 frame·s⁻¹。但比传统 SSD 算法检测速度低,略高于 DSSD 算法与 MFSSD 算法,这是由于本文改进算法模型采用不同尺度高低层特征融合方式导致模型复杂度较 SSD 算法高。但是目前安检机对管制物品的检测速率要求为 10 frame·s⁻¹[19],故本文算法的检测速度满足实际安检需求。

为验证用 ResNet34 替换 VGG16 和采用不同尺度高低层特征融合方式两种方法对原始 SSD 算法的改进是否有效,在 SDCK 管制刀具数据集上对本文改进算法的两个步骤进行分步测试,测试每个步骤的 mAP 如表 7 所示,其中原始 SSD 算法的 mAP 为 89.8。

表 7 改进算法分步测试 mAP

Table 7 Improved algorithm testmAP step by step

Base net	Function module	mAP / %
VGG16	—	89.8
ResNet34	—	90.8
VGG16	✓	90.5
ResNet34	✓	92.6

由表 7 可以看出利用残差网络 ResNet34 代替原 VGG16 作为 SSD 的基础网络,构建残差 SSD 检测器模型,并且本文改进算法对基础分类网络 ResNet34 后三层作相同卷积处理后,与原来特征层两两进行轻量级特征融合,得到新的三层特征图,使得基础网络对小目标提取的特征信息更加丰富,再采用跳跃连接的方式将高层特征图和低层特征图进行多尺度特征融合,增强了检测器模型对小目标刀具的特征提取能力,得到的检测器模型性能更佳。

本文改进的 SSD 综合提高了对安检图片中管制刀具小目标的检测能力,且算法精度高,鲁棒性好。实验中部分检测结果如图 6 所示。

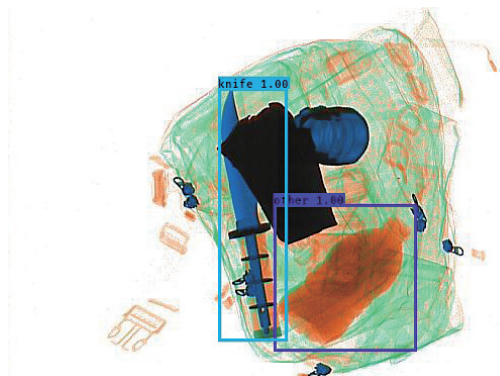


图 6 实验检测结果

Fig. 6 Experimental results

5 结 论

为进一步提高安检管制刀具中小目标的检测精度,改进原始 SSD 算法,完善了算法性能。采用 ResNet34 网络代替原 VGG16 作为 SSD 的基础网络,构建残差 SSD 模型,使得基础网络对小目标提取的特征信息更加丰富;再将 SSD-ResNet34 网络部分扩展层作反卷积形成新的高层网络,设计新高层网络与新低层网络跳跃连接的融合模块;改进后的模型利用多尺度多样化的特征图进行分类和回归任务,实验结果表明所提方法有效提高了对小目标的检测精度,且实时性好,能够满足实际安检需求。在下一步的研究中,将对数据集进行再度扩增,改进生成框预测阶段,减小生成框预测阶段的计算量,采用较小的卷积核和更窄的网络结构降低网络复杂度,提高算法的检测速度。

参 考 文 献

- [1] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [2] Zhang H Y, Wang S N, Hu W B. Improved method for estimating number of people based on convolution neural network [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121503.
张红颖, 王赛男, 胡文博. 改进的基于卷积神经网络的人数估计方法[J]. 激光与光电子学进展, 2018, 55(12): 121503.
- [3] Girshick R, Donahue J, Drell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of 2014 IEEE International Conference Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [4] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks [EB/OL]. (2016-05-20) [2020-08-15]. <https://arxiv.org/abs/1605.06409>.
- [7] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [8] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [9] Jeong J, Park H, Kwak N. Enhancement of SSD by concatenating feature maps for object detection [EB/OL]. (2017-05-26) [2020-08-15]. <https://arxiv.org/abs/1705.09587>.
- [10] Fu C Y, Liu W, Ranga A, et al. DSSD: deconvolutional single shot detector [EB/OL]. (2017-01-23) [2020-08-15]. <https://arxiv.org/abs/1701.06659>.
- [11] Leng J X, Liu Y. An enhanced SSD with feature fusion and visual reasoning for object detection [J]. Neural Computing and Applications, 2019, 31(10): 6549-6558.
- [12] Simonyan K, Zisserman A. Very deep convolutional

- networks for large-scale image recognition[EB/OL]. (2015-04-10) [2018-11-25]. <https://arxiv.org/abs/1409.1556v6>.
- [13] Jia Y Q, Shelhamer E, Donahue J, et al. Caffe: convolutional architecture for fast feature embedding [C] // Proceedings of the ACM International Conference on Multimedia-MM'14, November 3-7, 2014. Orlando, Florida, USA. New York: ACM Press, 2014: 675-678.
- [14] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [15] Cao G M, Xie X M, Yang W Z, et al. Feature-fused SSD: fast detection for small objects[J]. Proceedings of SPIE, 2018, 10615: 106151E.
- [16] Zhao Y N, Wu L M, Chen Q. Small object detection algorithm based on multi-scale fusion SSD [J]. Computer Engineering, 2020, 46(1): 247-254.
- 赵亚男, 吴黎明, 陈琦. 基于多尺度融合 SSD 的小目标检测算法 [J]. 计算机工程, 2020, 46(1): 247-254.
- [17] Erhan D, Szegedy C, Toshev A, et al. Scalable object detection using deep neural networks[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 2155-2162.
- [18] Szegedy C, Reed S, Erhan D, et al. Scalable, high-quality object detection [EB/OL]. (2014-12-03) [2020-08-15]. <https://arxiv.org/abs/1412.1441>.
- [19] Ji X L, Wu J, Yi J B, et al. Automatic detection algorithm for controlled items based on deep learning[J]. Laser & Optoelectronics Progress, 2019, 56(18): 180402.
- 吉祥凌, 吴军, 易见兵, 等. 基于深度学习的管制物品自动检测算法研究 [J]. 激光与光电子学进展, 2019, 56(18): 180402.