中国光学 Chinese Optics 核心期刊

结合空洞卷积和迁移学习改进YOLOv4的X光安检危险品检测

吴海滨 魏喜盈 刘美红 王爱丽 刘赫 岩堀祐之

Improved YOLOv4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning

WU Hai-bin, WEI Xi-ying, LIU Mei-hong, WANG Ai-li, LIU He, IWAHORI Yu-ji

引用本文:
吴海滨, 魏喜盈, 刘美红, 王爱丽, 刘赫, 岩堀祐之. 结合空洞卷积和迁移学习改进YOLOv4的X光安检危险品检测[J]. 中国光学, 2021, 14(6): 1417–1425. doi: 10.37188/CO.2021–0078
WU Hai-bin, WEI Xi-ying, LIU Mei-hong, WANG Ai-li, LIU He, IWAHORI Yu-ji. Improved YOLOv4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning[J]. *Chinese Optics*, 2021, 14(6): 1417-1425. doi: 10.37188/CO.2021-0078

在线阅读 View online: https://doi.org/10.37188/CO.2021–0078

文章编号　2095-1531(2021)06-1417-09

结合空洞卷积和迁移学习改进 YOLOv4 的 X 光安检危险品检测

吴海滨[1],魏喜盈[1],刘美红[1],王爱丽[1*],刘　赫[1],岩堀祐之[2]
(1. 哈尔滨理工大学 黑龙江省激光光谱技术及应用重点实验室,黑龙江 哈尔滨 150080;
2. 中部大学 计算机科学学院,爱知 春日井 487-8501)

摘要: 由于 X 光安检图像存在背景复杂,重叠遮挡现象严重,危险品摆放方式、形状差异较大等问题,导致检测难度较高。针对上述问题,本文在 YOLOv4 的基础上,结合空洞卷积对其网络结构进行改进,加入空洞空间金字塔池化 (Atrous Space Pyramid Pooling, ASPP) 模型,以此增大感受野,聚合多尺度上下文信息。然后,通过 K-means 聚类方法生成更适合 X 光安检危险品检测的初始候选框。其中,模型训练时采用余弦退火优化学习率,进一步加速模型收敛,提高模型检测精度。实验结果表明,本文提出的 ASPP-YOLOv4 检测算法在 SIXRay 数据集上的 mAP 达到 85.23%。该方法能有效减少 X 光安检图像中危险品的误检率,提高小目标危险品的检测能力。

关　键　词: X 光安检图像;YOLOv4;空洞卷积;空间金字塔池化;余弦退火
中图分类号: TP391.4;TH691.9　　**文献标志码:** A　　**doi:** 10.37188/CO.2021-0078

Improved YOLOv4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning

WU Hai-bin[1], WEI Xi-ying[1], LIU Mei-hong[1], WANG Ai-li[1*], LIU He[1], IWAHORI Yu-ji[2]
(1. *Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application, Harbin University of Science and Technology, Harbin* 150080, *China*;
2. *Department of Computer Science, Chubu University, Kasugai* 487-8501, *Japan*)
* *Corresponding author, E-mail: aili925@hrbust.edu.cn*

Abstract: In response to the complex backgrounds of X-ray security images, serious overlap and occlusion phenomena, and the large differences in the placement and shape of dangerous goods, this paper improves the network structure of YOLOv4 for dangerous objects detection by combining atrous convolution with the Atrous Space Pyramid Pooling (ASPP) model to increase receptive field and aggregate multi-scale context information. Then, the K-means clustering method is used to generate an initial candidate frame that is more suitable for dangerous goods detection in X-ray inspection images. Cosine annealing is used to optimize the learning rate in model training to further accelerate model convergence and improve model detection accuracy. The experimental results show that the proposed ASPP-YOLOv4 in this paper can obtain an mAP of

收稿日期: 2021-04-13; **修订日期:** 2021-05-11
基金项目: 国家自然基金科学基金(No. 61671190, No. 61801149); JSPS 科学基金(No. #20K11873)
Supported by National Natural Science Foundation of China (No. 61671190, No. 61801149); Japan Society for the Promotion of Science (No. #20K11873)

85.23% on the SIXRay dataset. The model can effectively reduce the false detection rate of dangerous goods in X-ray security images and improve the detection ability of small targets.

Key words: X-ray security images; YOLOv4; atrous convolution; spatial pyramid pooling; cosine annealing

1 引 言

随着社会的不断进步，公共交通在生产生活中所占据的位置日益重要。目前，安全检查广泛应用于一些政府机关办公场所、交通枢纽、军事基地、科研及保密场所、航天基地等场所，如果在这些场所发生意外，将会导致人员伤亡和经济财产损失，并带来极其恶劣的社会影响。为了减少此类恶性事件的发生，我国对安防领域的投入也逐渐增多，安检级别也越来越高。安检门和安检机是目前应用最广泛的安检设备。安检门，又称为金属探测门，主要是检查人身上携带的金属品，如枪支、械具等。安检机则主要是对人随身携带的行李、包裹等进行检查，通过 X 光安检通道对行李进行透视性扫描，安检员进而通过影像判断是否含有违禁品、危险品等。越来越密集的交通网络以及巨大的客流量使得我国的安检工作量激增，安全检查的工作面临严峻的挑战。

近年来，深度学习在目标检测等任务中发挥了至关重要的作用[1-2]。利用深度卷积神经网络能自动提取 X 光危险品特征的优势，有效地解决了物品遮挡、背景杂乱等问题，提高了危险品的检测性能。2017 年，Mery 等在 GDXray 数据集测试了 10 余种安检图像识别方法，它们包括基于 BoW、稀疏 KNN、深度学习等，这也是深度学习首次应用到 X 光安检领域[3]。2018 年，Akcay 提出了一种基于深度学习的行李检查方法。首先，标记图像中的每个对象，然后将图像的位置和边界框指定给具有卷积神经网络的区域，最后检测并识别出危险品[4]。Morris等人提出了 Passenger Baggage Object Database，应用于危险品分类，并且在 VGG19、Xception、InceptionV3 模型上进行测试，得到最佳模型的 ROC 曲线为 0.95[5]。Akcay 等采用迁移学习 (Transfer Learning) 的方法，使用深度卷积神经网络解决 X 光行李安全筛选背景下的图像分类问题[6]。

2019 年，Akcay 提出了 GANomaly 模型，利用生成对抗网络训练不包含危险品的正常样本，将包含危险品的少数样本作为异常样本，通过判别器判断其与网络学习到的潜在变量之间的差异，从而判断是否为异常样本[7]。Reagan 等人利用 VGG-16 网络在 ImageNet 数据上训练了 1 000 个不同的类，再将该模型学习到的内容作为初始化参数迁移到危险品的分类任务上，大大改善了分类精度[8]。

2020 年，唐浩漾等人提出在 SSD 目标检测网络上构建特征金字塔，融合危险品的底层边缘形状信息特征和深层语义特征，并使用可行变卷积提高网络危险品检测能力[9]。张友康等人提出了非对称卷积多视野检测网络，该网络使用小卷积非对称网络、多尺度特征图融合策略以及空洞多视野卷积模块，有助于提高背景干扰下的 X 光安检图像危险品识别[10]。郭守向等人构建了基于 YOLO (You Only Look Once) 的 X 光图像违禁品检测网络，其主干网络 Darknet-C 由辅助骨干网络各个特征层和引导主干网络对应的上一层级进行特征级联后向下一层传播，有效提升了检测小目标的检测效果[11]。Zhu 等人提出了一种用于 X 光安检图像的数据增强方法，丰富安检图像的多样性。首先，使用改进的自我注意生成对抗网络 (Self-Attention Generative Adversarial Networks, SAGAN) 生成了新的危险品图像。其次，提出了一种基于循环生成对抗网络 (Cycle Generative Adversarial Networks, CycleGAN) 的方法，将物品的自然图像转换为 X 光图像[12]。

本文在 YOLOv4 的基础上，结合空洞卷积对其网络结构进行改进，加入空洞空间金字塔池化模型 (Atrous Spatial Pyramid Pooling, ASPP)，称为 ASPP-YOLOv4 模型，以此增大感受野，聚合多尺度上下文信息，同时利用迁移学习训练网络和余弦退火优化学习率，大大提高了对 X 光安检图像中危险品的检测精度。

2 ASPP-YOLOv4 检测模型

本文构建的 ASPP-YOLOv4 模型包括模型预训练、网络结构、锚框选择 3 个部分，如图 1 所示。

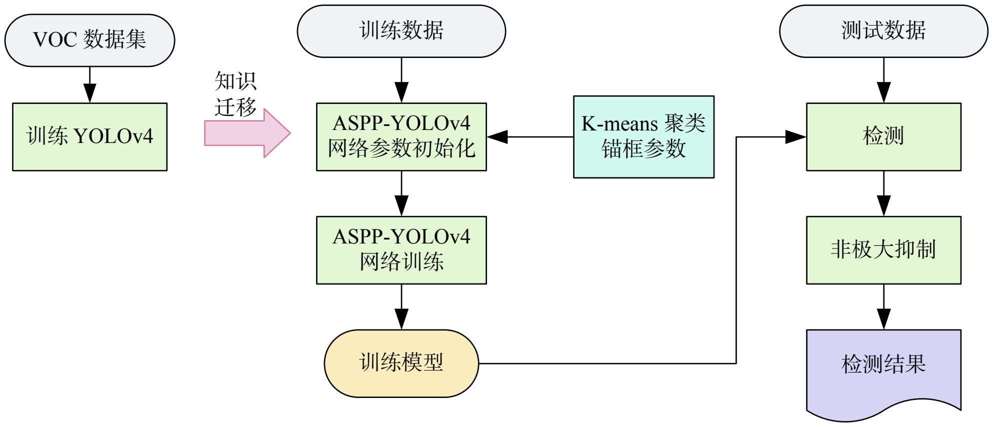


图 1　ASPP-YOLOv4 模型设计

Fig. 1　ASPP-YOLOv4 model design

(1)引入迁移学习, 利用公开的 VOC 数据集对 YOLOv4 网络进行预训练, 将训练好的参数迁移到本文构建的 ASPP-YOLOv4 模型进行微调。与正常卷积神将网络训练中的随机初始化不同, 将预先训练的网络权重值作为初始化参数加载, 可以有效加速收敛并提高检测性能。

(2)采用具有高性能和强稳定性的 YOLOv4 作为基本检测模型, 结合上下文信息对于检测目标的重要性, 将原网络中的空间金字塔池化改进为空洞空间金字塔池化, 构建 ASPP-YOLOv4 检测网络。通过使用不同扩张率的空洞卷积实现不同尺度的最大值池化, 在增大感受野的同时有效融合多尺度的上下文信息, 增强模型对危险品尺度变化的鲁棒性。

(3)根据危险品目标的尺度分布特点重新设计了初始锚框。使用 K-means 聚类算法对危险品目标的尺度进行聚类的结果作为初始锚框, 使得网络候选框的初始化尺寸更加贴合危险品目标。另外, 在训练过程中采用学习率余弦退火衰减策略, 实现模型训练过程的优化。

2.1　YOLOv4

YOLO[13-14] 网络是一种基于回归的目标检测网络, 其可以通过充分提取被检目标的特征, 直接在一个网络内同时完成目标的分类和定位任务, 真正实现了一种简单且高效的端到端的设计思路, 成为 one-stage 检测模型的重要代表。

YOLOv4[15-16] 算法是在 YOLOv3 的基础上, 从数据增强、主干网络、网络训练、损失函数等方面进行改进优化, 虽然其网络结构更复杂, 但有更高的精度和更快的速度, 在检测小目标和被遮挡目标方面有着显著的提高。图 2 为 YOLOv4 的网络结构。

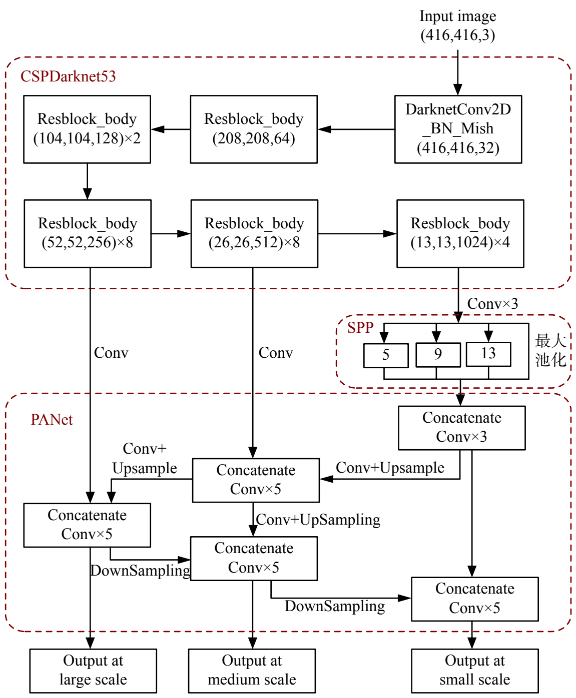


图 2　YOLOv4 网络结构

Fig. 2　Network structure of YOLOv4

YOLOv4 算法主要包括 4 个步骤:

(1) 将输入图片分成$S \times S$个网格, 物品中心落在网格中心, 检测该网格中的物体;

(2) 每个网格都生成B个预测框;

(3) 每一个预测框均针对C个类别进行概率预测;

(4) 使用非极大值抑制算法对所有预测框进行筛选, 得到最终包含目标类别和位置坐标的检测结果。

2.2 空洞空间金字塔池化模型

在 X 光安检图像检测任务中，容易受到遮挡、拍摄角度、复杂背景等因素的影响，导致其外观信息获取不足，影响其整体信息的收集，因此在检测危险品时，仅通过目标表征进行判断的准确率较低。上下文信息具有描述目标内在联系的特点，将其加入至目标检测任务中，能有减少上述影响，提高检测的准确度。空洞卷积在不使用池化层的情况下，能够增大感受野，使每个卷积输出都包含较大范围的信息，进而提高网络性能。其核心是向卷积层引入一个新参数——扩张率 (dilation rate)，该参数定义了卷积核处理的间隔数量。空洞卷积核和感受野大小计算如式 (1) 和式 (2) 所示：

$$f_k' = f_k + (f_k - 1) * (r - 1) \quad , \tag{1}$$

$$R_m = R_{m+1} + (f_k' - 1) * \prod_{i=1}^{m-1} s_i \quad , \tag{2}$$

其中，f_k表示原始卷积核的大小，f_k'表示空洞卷积核大小，r表示卷积核的扩张率，R_m表示经空洞卷积后第m层感受野大小，S_i表示第i层的步长，“*”表示卷积。

YOLOv4 中使用 SPP 模块提取不同感受野的信息，但没有充分体现全局信息和局部信息的语法关系。本文设计的 ASPP 引入不同扩张率的空洞卷积操作，实现 SPP 中的池化操作，并将其与全局平均池化并联，组成一个新的特征金字塔模型，以此聚合多尺度上下文信息，增强模型识别不同尺寸同一物体的能力。

结合 ASPP 改进的 YOLOv4 框架如图 3 所示，其中 ASPP 采用扩张率 r 分别为 1、3、5 且卷积核为3×3的空洞卷积，在前一层的局部特征上关联到更广的视野，防止小目标特征在信息传递时丢失。

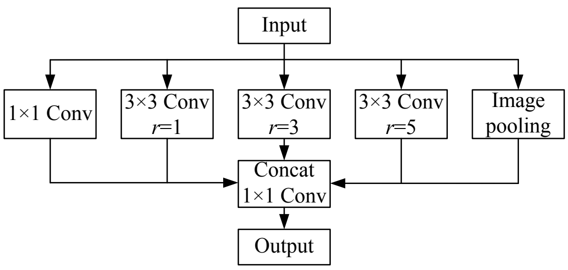


图 3 结合 ASPP 改进的 YOLOv4 框架

Fig. 3 Improved YOLOv4 framework combined with ASPP

图 3 中，ASPP 将输入的特征同时输出到 5 个支路中，第 1 个支路是一个 1×1 的卷积，目的是保持原始的感受野；第 2~4 支路是不同扩张率的空洞卷积，并且包含 BN 层，目的是对特征进行提取，获得不同的感受野，最后的特征图能够融合不同尺度的信息；第 5 个支路的目的是获取全局特征，将特征做全局平均化池化后，经过 1×1 卷积和归一化处理，再进行上采样。本文模型通过 ASPP，可以学习到不同尺寸危险品的特征，降低危险品的漏检率。

通过特征融合后，YOLOv4 获得了 3 个特征层的预测结果，尺寸大小分别为 (13, 13, 30), (26, 26, 30), (52, 52, 30) 的数据，分别对应网格上 3 个预测框的位置，其中每一个网格点负责一个区域的检测。其中维度 30 代表了 3 个 box 的详细信息，每个 box 长度为 10。这 10 个维度的信息为得到的预测结果：$(x\_offset, y\_offset, h, w, conf, conf_1, \cdots, conf_5)$。其中$x_offset$表示网格左上角相对于$x$轴的偏移量，$y_offset$代表网格左上角相对于$y$轴的偏移量，$h$代表预测框的高，$w$代表预测框的宽，$conf$代表置信度，$conf_i(1 \leqslant i \leqslant 5)$表示分类结果。最后，取出每一类得分大于阈值的框和得分进行排序，利用框的位置和得分进行非极大抑制，最终得出概率最大的边界框。

2.3 利用迁移学习训练网络

在深度神经网络中，前面的卷积层一般学习具有通用性的浅层特征，而后面的卷积层学习的是对当前训练目标更具有针对性的、更高级的抽象特征。先冻结部分网络层可以加快训练速度，也可以在训练初期防止权值被破坏。导致模型泛化能力差的另一个原因是目标函数存在局部最小值，而迁移学习能有效避免这个问题。

考虑到存在大量可见光图像，而 X 光危险品图像过少，采用基于参数迁移的学习方法训练 ASPP-YOLOv4 网络。迁移策略如下：

(1)选取公开的 VOC 数据集为预训练数据集，该数据集有 20 类标注目标，共计 16 万张图像，用 VOC 数据集为源域在 YOLOv4 网络上学习图片底层特征；

(2)取训练检测效果较好的 YOLOv4 源域网络，去掉后面全连接层和分类层，作为 ASPP-

YOLOv4 骨干网络的预训练权重, 其他网络部分采用权重随机初始化的方式;

(3)冻结主干网络, 对不冻结的网络层进行微调, 参与反向传播中的梯度更新;

(4)放开冻结层, 对全部层利用反向传播算法进行训练, 得到最终合适的参数矩阵和偏置向量。

2.4　基于 K-means 聚类的锚框选择

采用 K-means 聚类的方法代替原始锚框, 使得网络候选框的初始化尺寸更加贴合危险品目标, 进一步提高网络的检测效果。K-means 是一种无监督学习算法, 它将数据集划分成多个簇, 通过计算每个样本的相似性度量, 并将其分配给与之距离最近的簇中心, 每次迭代都计算每个簇中所有样本的均值, 不断更新簇中心, 直到簇中心不再变化。

本文处理的 X 光危险品的真实框由左上定点和右下定点表示, 即($x_{\min}, y_{\min}, x_{\max}, y_{\max}$)。由于数据集中图片大小不同, 首先对使用图片的长宽对 box 的宽和高按式 (3) 做归一化处理。标准的 K-means 使用欧式距离作为度量, 但本文检测方法中只需要关心锚框与真实框的交并比 (Intersection Over Union, IOU), 不关心真实框的大小, 因此, 使用 IOU 作为度量更加合适。假设$anthor=(w_a, h_a)$, $box=(w_b, h_b)$, 则二者的 IOU 如式 (4) 所示。IOU 值越大, 表示两个框距离越近。最终度量如式 (5) 所示。

$$w_b = \frac{w_{box}}{w_{img}}, h_b = \frac{h_{box}}{h_{img}} \quad , \tag{3}$$

$$\begin{aligned} &\mathrm{IOU}(box, anchor) = \\ &\frac{intersection(box, anchor)}{union(box, anchor) - intersection(box, ancher)} = \\ &\frac{\min(w_a, w_b)\cdot\min(h_a, h_b)}{w_a h_a + w_b h_b - \min(w_a, w_b)\cdot\min(h_a, h_b)} \quad , \end{aligned} \tag{4}$$

$$d(box, anchor) = 1 - \mathrm{IOU}(box, anchor) \quad . \tag{5}$$

本文首先选取 9 个框作为初始的锚框, 再使用 IOU 度量, 将每个框分配给其距离最近的锚框形成新的簇, 并在每次迭代后计算每个簇中所有框的均值, 更新锚框, 直到锚框不再变化。最终得到的锚框如表 1 所示。

表 1　锚框计算结果

Tab. 1　Calculation results of the anchor

特征图	感受野	*anchor*
13×13	大	(124×111) (171×61) (200×151)
26×26	中	(75×34) (82×188) (93×75)
52×52	小	(24×78) (50×67) (62×111)

2.5　余弦退火优化学习率

ASPP-YOLOv4 的学习率采用了余弦退火 (Cosine annealing) 优化算法, 它与传统的学习率不同, 随着 epoch 的增加, 学习率先模拟余弦函数快速下降, 再线性上升, 不断重复该过程。传统的训练过程中学习率逐渐减小, 模型逐渐寻找到局部最优点, 模型会快速地往局部最优点移动, 随着学习率的减小, 模型最终收敛到一个全局最优点。而余弦退火的学习率急速下降, 所以模型会迅速踏入局部最优点, 并保存局部最优点的模型, 然后开启热重启 (Warm restart), 学习率重新恢复到一个较大值, 逃离当前的局部最优点, 并寻找新的最优点, 周期性重复上述过程。学习率余弦退火衰减具体策略如表 2 所示。

表 2　余弦退火衰减过程

Tab. 2　Cosine annealing decay process

算法: 余弦退火衰减算法

输入: 训练epoch E_p、训练批次B_s、预热期w_epoch、预先设置学习率η_{base}、最大学习率$\eta_{\max}$、最小学习率$\eta_{\min}$、训练样本数S_c;

输出: 当前训练学习率η_t

步骤:

(1) 初始化总步长$Steps_{total} = \left(E_p \times S_c\right)/B_s$
预热步长$Steps_{warmup} = (w \times S_c)/B_s$

(2) Repeat:

在每次重启之后执行:

更新当前执行的步数$step_{global}$, 并记录当前学习率

更新学习率

if $Steps_{global} < Steps_{warmup}$:

根据$\eta_t = \left((\eta_{base} - \eta_{warmup})/Steps_{warmup}\right) \times Steps_{global} + \eta_{warmup}$
计算线性增长的学习率η_{warmup}

else:

根据$\eta_t = \dfrac{1}{2} \times \eta_{base} \times \cos\left(1 + \left(\pi \times \dfrac{(Steps_{global} - Steps_{warmup})}{Steps_{total} - Steps_{warmup}}\right)\right)$
计算余弦退火的学习率

$\eta_t = \min(\eta_t, \eta_{\min})$

其中重启并不是从头开始，而是通过增加学习率来模拟，另外定义了一个初始学习率。由于模型在刚开始训练时的权重是随机初始化的，如果此时就直接采用较大的学习率，容易造成模型的震荡。采用 warmup 预热学习率可以避免上述问题，这是因为如果在刚开始训练的几个 epoch 内采用预热的小学习率下，模型可以慢慢趋于稳定，这时再使用预设的学习率进行训练，可获得更快的收敛速度。

3 实验结果及分析

本文实验均在 Windows 操作系统下实现，采用 Python 语言和 Keras 库编码，NVIDIA RTX2070 显卡 GPU 加速，CPU 为 32 GB Intel® Core™ i7-9750H CPU @ 2.60 GHz。数据集采用的是 SIXRay 数据集[17]中所有正类样本添加位置标签生成的数据集，手动框出危险品的位置并标注出类别，最终生成每幅图片的标签文件。将数据集划分为两部分，80% 样本作为训练集，20% 样本作为测试集。选择余弦退火学习率算法作为梯度优化算法。

训练超参数配置如表 3 所示。为了验证本文算法的有效性，选取目前主流目标检测模型，包括 YOLOv3[18]、M2Det[19]、SSD[20]、YOLOv4 做对比。表 4 列出了以上几种方法对 5 种危险品的检测结果。

表 3 训练超参数设计

Tab. 3 Design of the training hyperparameters

状态	名称	参数
冻结主干网络	batch_size	8
	epoch	50
	最大学习率	1e-3
	最小学习率	1e-6
	Warmup_epoch	10
解冻主干网络	batch_size	2
	epoch	50
	最大学习率	1e-4
	最小学习率	1e-6
	Warmup_epoch	10

表 4 不同模型的 AP 比较

Tab. 4 Comparison of AP for different networks (%)

方法	AP					mAP
	Gun	Knife	Wrench	Pliers	Scissors	
YOLOv3	93.18	78.00	68.55	79.69	76.97	79.28
M2Det	95.49	75.70	70.17	83.00	82.96	81.47
SSD	94.91	77.87	74.82	84.51	82.69	82.96
YOLOv4	94.40	81.69	77.38	84.50	77.55	83.11
ASPP-YOLOv4	95.78	81.39	77.84	87.36	83.76	85.23

在目标检测模型训练时，训练和验证过程中每一个 Epoch 的损耗值下降曲线如图 4 所示。从图中可以看出，在前 50 个 Epoch 训练过程中，随着迭代次数和训练时间的增加，Train-Loss 和 Validation-Loss 均逐渐下降，最终趋于平稳，训练过程中设置了早停策略，得到最优检测模型。

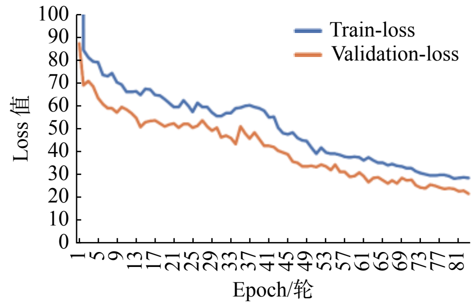


图 4 训练过程中的 Loss 下降曲线

Fig. 4 Loss decline curves during training process

对于 Gun 来说，数量相对较多，特征明显，所有网络对其检测精度较高；由于 Wrench 的类内差异大、相互重叠，所有网络对它们的检测精度均偏低；Scissors 的数量少，目标小，易被覆盖，漏检率较高。ASPP-YOLOv4 的多类别平均精度 (mean Average Precision, mAP) 为最高，达到 85.23%，相比于 YOLOv3、M2Det、SSD、YOLOv4 分别提升了 5.95%、3.76%、2.27% 和 2.12%。除了 Knife，本文的检测精度均最高。

与 YOLOv4 相比，Gun、Wrench、Pliers 的平均精度 (Average Precision, AP) 分别提升了 1.38%、0.46%、2.86%，Scissors 提升明显，为 6.21%。对于 Knife 来说，本文方法相对于 YOLOv3 提升了 3.39%，比 M2Det 提升了 5.69%，比 SSD 提升了 3.52%，与 YOLOv4 相比，仅下降了 0.3%，是可以接受的。

图 5(彩图见期刊电子版) 用柱状图比较 5 种方法在每类样本上的 AP, 可以直观地看出, 本文方法具有良好的检测效果, 性能最优。

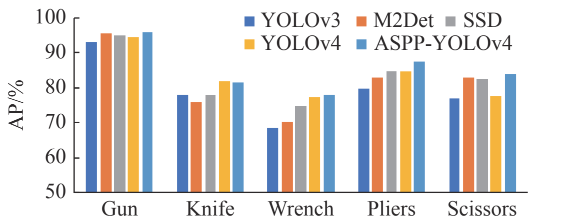


图 5　几种方法对各类危险品的检测结果

Fig. 5　Detection results of dangerous goods detected by different algorithms

表 5 列出了 ASPP-YOLOv4 对 SIXRay 数据集上不同类别危险品的检测性能。从表 5 可以看出, 针对各类样本, 模型的 AP 较好, 但召回率 (Recall) 相对较低, 这是因为安检图像目标较小, 重叠较多, 更容易造成漏检而不是误检。尤其是对于 Knife 这类物品, 摆放的方式、物品大小形状等差异较大, 漏检率相对较高。而本模型的准确率 (Precision) 保持在较好的水平, 误检率较低。

表 5　ASPP-YOLOv4 的性能分析

Tab. 5　The performance of ASPP-YOLOv4

类别	AP	Precision	Recall	F1-Measure
Gun	95.78%	98.44%	85.32%	0.91
Knife	81.39%	91.48%	67.40%	0.78
Wrench	77.84%	81.61%	71.05%	0.76
Pliers	87.36%	93.15%	75.79%	0.84
Scissors	83.76%	86.28%	76.23%	0.81

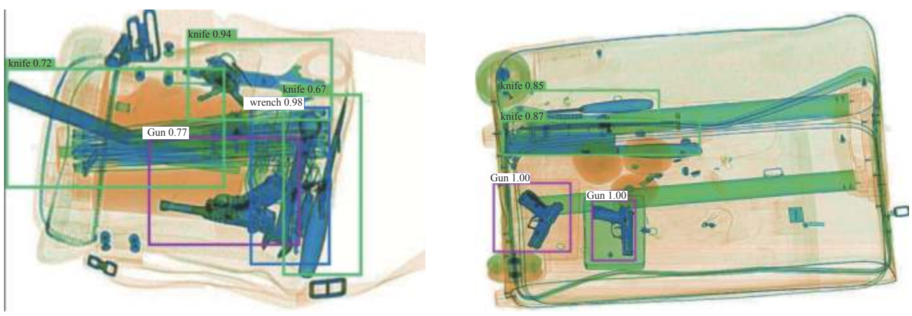
为了进一步证明本文提出的 ASPP-YOLOv4 算法的有效性, 表 6 取各类目标的平均值, 作为整体性能评估指标, 与原始 YOLOv4 进行对比。从表 6 中可以看出, ASPP-YOLOv4 相比 YOLOv4, mAP 提高了 2.12%, 召回率提高了 2.16%。表明本文方法能更好地与目标特点匹配, 目标定位更准确, 同时空洞卷积的引入提高了模型识别同一类物品的能力, 所以能有效提升目标准确率和召回率。

表 6　YOLOv4 改进前后检测性能对比

Tab. 6　Comparison of YOLOv4 performance before and after improvement

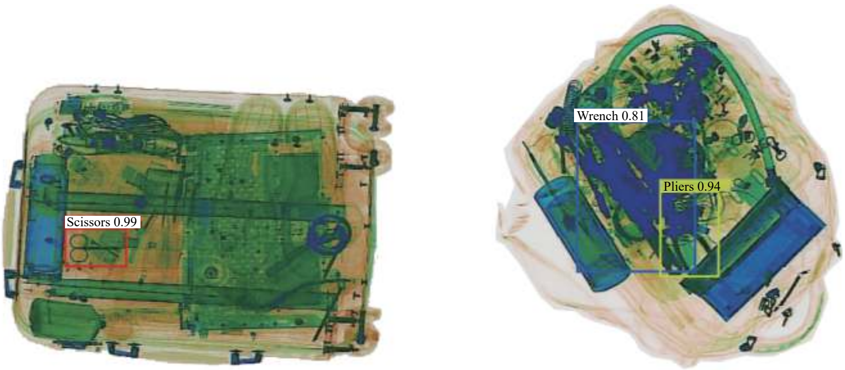
方法	mAP	Precision	Recall	F1-Measure
YOLOv4	83.11%	90.35%	73.00%	0.80
ASPP-YOLOv4	85.23%	90.20%	75.16%	0.82

ASPP-YOLOv4 模型在测试集上的目标检测效果如图 6(彩图见期刊电子版) 所示。其中图 6(a) 是多目标图像, 可以看出 ASPP-YOLOv4 模型能较为准确地检测出每一个目标, 其中绿色检测框表示刀具, 同一张图片中的 3 个刀具具有不同的形状和大小, 但是本文模型可以较好地识别, 置信度分别为 0.94、0.72、0.67; 紫色检测框表示枪支, 置信度为 0.77; 蓝色检测框表示扳手, 由于轮廓相对清晰, 置信度达到 0.98。图 6(b) 是摆放差异图像, 两把刀具的摆放方式有较大的差异, 但背景较为简单, 模型对两支枪的置信度均为 1.00, 对刀具的置信度分别为 0.85 和 0.87。图 6(c) 是小目标图像, 红色检测框表示剪刀, ASPP-YOLOv4 对其有较好的识别精度, 置信度为 0.99。图 6(d) 是重叠图像, 背景杂乱且颜色相近, ASPP-YOLOv4 模型可以获得更大的感受野, 较好地融合上下文信息, 进而能从重叠的 X 光图像中得到更加准确的目标特征, 对扳手的置信度为 0.81, 黄色检测框表示钳子, 其置信度为 0.94。



(a) 多目标图像
(a) Multiple targets

(b) 摆放差异图像
(b) Different positions



(c) 小目标图像
(c) Small targets

(d) 重叠图像
(d) Overlapping targets

图 6　危险品检测效果

Fig. 6　Detection results for dangerous goods

4　结　论

本文对 YOLOv4 进行改进，实现了 X 光安检图像危险品检测，该模型结合空洞卷积改进空间金字塔池化，在增大感受野的同时有效融合多尺度的上下文信息。同时利用迁移学习训练网络和余弦退火优化学习率，提高检测精度的同时，加速模型收敛，减少训练时间。最后以 K-means 聚类的方法设计锚框，使其更加适合危险品的定位。实现结果显示，在 SIXRay 数据集上的 mAP 达到 85.23%，充分表明 ASPP-YOLOv4 模型对于 X 光安检危险品图像的检测结果优于其他方法，能有效解决物体遮挡问题，危险品的分类与定位效果均有提升。

参考文献:

[1] 鞠默然，罗海波，刘广琦，等. 采用空间注意力机制的红外弱小目标检测网络[J]. 光学 精密工程，2021，29(4)：843-853.
JU M R, LUO H B, LIU G Q, *et al*.. Infrared dim and small target detection network based on spatial attention mechanism[J]. *Optics and Precision Engineering*, 2021, 29(4): 843-853. (in Chinese)

[2] 马立，巩笑天，欧阳航空. Tiny YOLOV3目标检测改进[J]. 光学 精密工程，2020，28(4)：988-995.
MA L, GONG X T, OUYANG H K. Improvement of Tiny YOLOV3 target detection[J]. *Optics and Precision Engineering*, 2020, 28(4): 988-995. (in Chinese)

[3] MERY D, SVEC E, ARIAS M, *et al*.. Modern computer vision techniques for X-ray testing in baggage inspection[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017, 47(4): 682-692.

[4] AYDIN I, KARAKOSE M, AKIN E. A new approach for baggage inspection by using deep convolutional neural networks[C]. *2018 International Conference on Artificial Intelligence and Data Processing* (*AIDP*), *IEEE*, 2018: 1-6.

[5] MORRIS T, CHIEN T, GOODMAN E. Convolutional neural networks for automatic threat detection in security X-ray images[C]. *2018 17th IEEE International Conference on Machine Learning and Applications* (*ICMLA*), *IEEE*, 2018: 285-292.

[6] AKCAY S, KUNDEGORSKI M E, WILLCOCKS C G, *et al*.. Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(9): 2203-2215.

[7] AKÇAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Skip-GANomaly: skip connected and adversarially trained encoder-decoder anomaly detection[C]. *Proceedings of 2019 International Joint Conference on Neural Networks* (*IJCNN*), *IEEE*, 2019: 1-8.

[8] GALVEZ R L, DADIOS E P, BANDALA A A, *et al*.. Threat object classification in X-ray images using transfer

learning[C]. *Proceedings of 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), IEEE*, 2018: 1-5.

[9] 唐浩漾, 王燕, 张小媛, 等. 基于特征金字塔的X光机危险品检测算法[J]. 西安邮电大学学报, 2020, 25(2): 58-63.
TANG H Y, WANG Y, ZHANG X Y, *et al*.. Dangerous goods detection algorithm by X-ray machine based on feature pyramid[J]. *Journal of Xi'an University of Posts and Telecommunications*, 2020, 25(2): 58-63. (in Chinese)

[10] 张友康, 苏志刚, 张海刚, 等. X光安检图像多尺度违禁品检测[J]. 信号处理, 2020, 36(7): 1096-1106.
ZHANG Y K, SU ZH G, ZHANG H G, *et al*.. Multi-scale prohibited item detection in X-ray security image[J]. *Journal of Signal Processing*, 2020, 36(7): 1096-1106. (in Chinese)

[11] 郭守向, 张良. Yolo-C: 基于单阶段网络的X光图像违禁品检测[J]. 激光与光电子学进展, 2021, 58(8): 0810003.
GUO SH X, ZHANG L. Yolo-C: one-stage network for prohibited items detection within X-ray images[J]. *Laser & Optoelectronics Progress*, 2021, 58(8): 0810003. (in Chinese)

[12] ZHU Y, ZHANG Y T, ZHANG H G, *et al*.. Data augmentation of X-ray images in baggage inspection based on generative adversarial networks[J]. *IEEE Access*, 2020, 8: 86536-86544.

[13] 陈科峻, 张叶. 基于YOLO-v3模型压缩的卫星图像船只实时检测[J]. 液晶与显示, 2020, 35(11): 1168-1176.
CHEN K J, ZHANG Y. Real-time ship detection in satellite images based on YOLO-v3 model compression[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(11): 1168-1176. (in Chinese)

[14] 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), IEEE*, 2016: 779-788.

[15] 刘杨帆, 曹立华, 李宁, 等. 基于YOLOv4的空间红外弱目标检测[J]. 液晶与显示, 2021, 36(4): 615-623.
LIU Y F, CAO L H, LI N, *et al*.. Detection of space infrared weak target based on YOLOv4[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(4): 615-623. (in Chinese)

[16] BOCHKOVSKIY A, WANG C Y, LIAO H Y M.YOLOv4: optimal speed and accuracy of object detection[J/OL]. arXiv: 2004.10934, 2020(2020-04-23). https://arxiv.org/abs/2004.10934.

[17] MIAO C J, XIE L X, WAN F, *et al*.. SIXray: A large-scale security inspection X-ray benchmark for prohibited item discovery in overlapping images[C]. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE*, 2019: 2114-2123.

[18] REDMON J, FARHADI A.YOLOv3: an incremental improvement[J]. arXiv e-prints arXiv: 1804.02767, 2018.

[19] ZHAO Q J, SHENG T, WANG Y T, *et al*.. M2Det: a single-shot object detector based on multi-level feature pyramid network[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 9259-9266.

[20] LIU W, ANGUELOV D, ERHAN D, *et al*.. SSD: single shot multibox detector[C]. *14th European Conference on Computer Vision (CVPR), Springer*, 2016: 21-37.

作者简介:

吴海滨(1977—), 男, 上海人, 博士, 教授(博导), 2002 年于哈尔滨工业大学获得硕士学位, 2008 年于哈尔滨理工大学获得博士学位, 现为哈尔滨理工大学测控技术与通信工程学院教授, 主要从事机器视觉、医学虚拟现实、深度学习图像分类研究。E-mail: woo@hrbust.edu.cn

王爱丽(1979—), 女, 天津人, 博士, 副教授, 硕士生导师, 2008 年于哈尔滨工业大学获得博士学位, 现为哈尔滨理工大学测控技术与通信工程学院副教授, 主要从事机器视觉、深度学习图像分类研究。E-mail: aili925@hrbust.edu.cn