

基于深度学习算法的包裹携带品自动化检测模型

摘 要：包裹携带品检测作为物流行业和安防行业的关键工作，承担着防止危险违禁品进入运输渠道的重要任务。“时间紧、任务重、人员少”始终是摆在安检工作面前的难题，而伴随着客运物流行业的高速发展，巨量的包裹已经远超人工可以处理的范围，给物流行业和安防行业带来了巨大的挑战。近些年，深度学习技术的突破使得高精度的目标检测技术得以广泛应用，并取得了骄人战绩，这使得将目标检测技术应用于安检工作成为可能。本文基于三种目标检测算法分别建立模型，对包含五类违禁品的 1406 张图像数据进行学习，在训练 100000 次之后，最优 mAP 稳定在 83.62%左右。使用多张 X 光透射图像进行测试，检测模型成功检测出了图像中的违禁物品。最后，对于模型部署提出了“端云结合”的部署方式，并对检测模型的实际应用和改进方向给出了建议。

关键词：物流安检；自动化；检测算法；深度学习；云计算

Automatic Detection Model of Parcels Based on Deep Learning Algorithms

Abstract: As the key work of logistics and security industry, the detection of package carriers undertakes the important task of preventing dangerous contraband from entering the transport channels. Tight time, heavy tasks and less personnel, which has always been a difficult problem in front of security inspection. With the continuous development of passenger logistics industry, the huge number of parcels has far exceeded the scope of manual processing, which has brought great challenges to the logistics industry and security industry. In recent years, the breakthrough of deep learning technology has made high-precision target detection technology widely used and achieved remarkable achievements, which makes it possible to apply target detection technology to security inspection work. Based on three target detection algorithms, this paper establishes models to learn 1406 image data containing five types of contraband, after training for 100,000 times, the optimal mAP was stable at about 83.62%. Using several X-Ray images to be tested, the detection model successfully detects the prohibited items in the image. At last, the deployment mode of "combination of the end and cloud" is proposed for model deployment, and the suggestions of the practical application and improvement direction of the detection model are proposed.

Keywords : logistics security inspection; automation; detection algorithm; deep learning; cloud computing

0 引言

2018 年天猫“双十一”购物节单日成交总额突破 2135 亿元,创历史新高,电商经济的飞速发展给物流行业带来巨大机遇的同时也带来了非同一般的挑战。为了确保人民群众的生命财产安全,对包裹进行细致而全面的检查是物流行业的重要义务,然而“时间紧、任务重、人员少”始终是摆在安检工作面前的一大难题,人眼长时间的工作会因为疲劳而使得误判率直线上升,这将给安检工作带来巨大隐患。

不仅是物流行业,客运行业的安检工作形势同样严峻。2017 年,我国运输服务旅客发送量 185.1 亿人,相较于物流运输行业,人员基数大,身份错综复杂是客运行业的特点,这对客运安检工作有着更高的要求。旅客人均安检时间少、待排查范围广,这使得有限的安检人员常常力不从心,如何显著地提高安检工作的效率和准确性,是摆在整个行业面前的巨大难题。

近些年,深度学习技术的突破使得高精度的目标检测技术得以广泛应用,这使得将目标检测技术应用于安检工作成为可能。在本文中,我们将基于三种主流目标检测算法设计包裹携带品自动化检测模型,然后对五类“限制品”的 X 光透射图像进行检测。

1 数据处理

1.1 数据分析

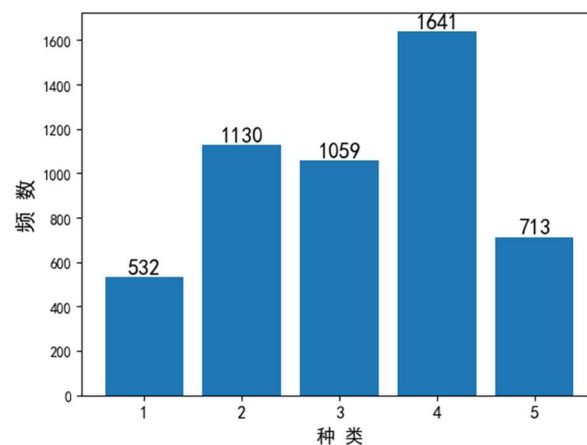
本次采用的源数据集共有 1406 张标注图像,其中包含五类“限制品”,如下表所示:

Table 1 限制品种类

编 号	名 称
1	铁壳打火机
2	黑钉打火机
3	刀 具
4	电池电容
5	剪 刀

其中,五类限制品出现频数统计如下:

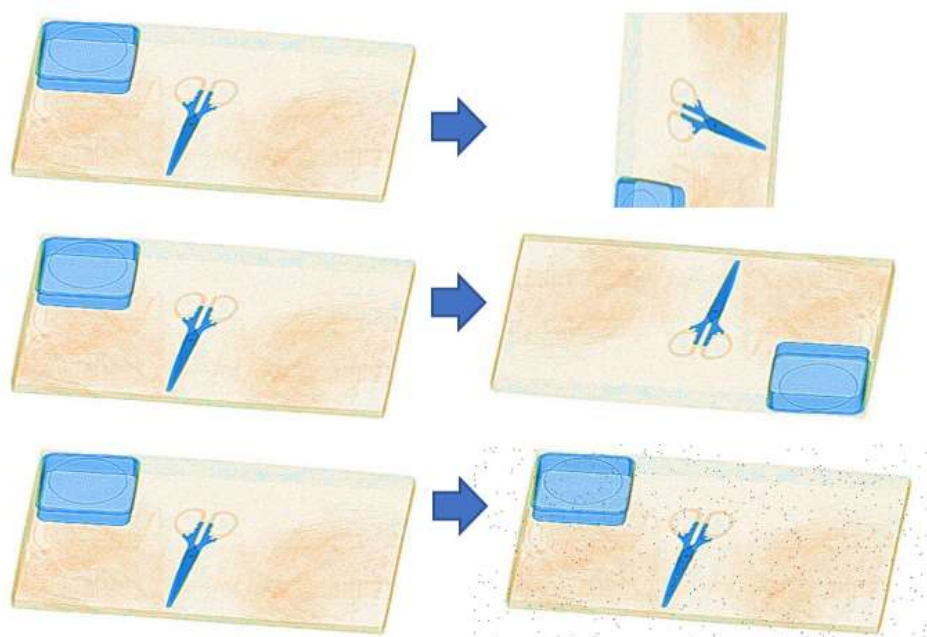
Figure 1 限制品出现频数条形图



1.2 数据增强

由于原数据实例较少,为了扩大数据样本量,提升数据空间的复杂度,首先对源数据进行增强处理,处理结果如图所示:

Figure 2 图像增强实例
(从上到下依次为对原始图像进行旋转、镜像、添加噪声)



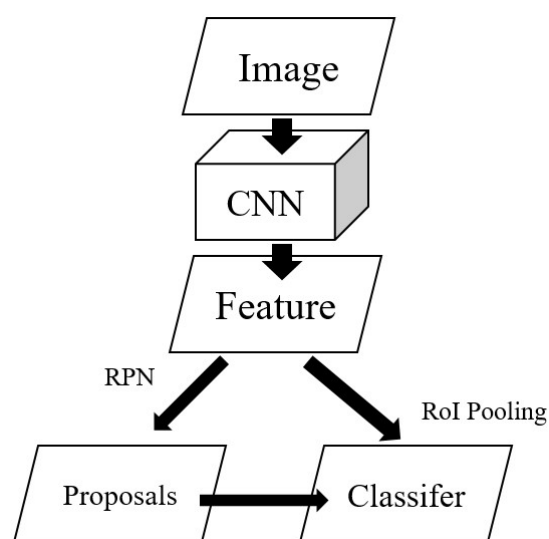
通过数据增强，将原有的 1406 张原图像扩张为 3000 张图像，使数据规模和数据特征的复杂程度得以显著提升。

2 基于 Faster R-CNN 算法的检测模型

2.1 Faster R-CNN 算法原理

Faster R-CNN 算法作为典型的“两阶段”目标检测算法，将目标检测任务分为两个过程——候选区域生成、特征提取及分类。“候选区域生成”阶段首先通过卷积网络提取图像特征，然后使用 RPN(Region Proposal Network) 网络产生若干候选区域。^{[1][2][3]}

Figure 3 Faster R-CNN 算法结构



在 RPN 生成若干候选区域之后，Faster R-CNN 再使用卷积网络和 ROI 池化进行特征提取和特征分类，作为“两阶段”算法的代表，Faster R-CNN 算法在 VOC 2007 数据集上获得了 78.8% 的平均精度。^{[1][2][3]}

Figure 4 RPN 网络

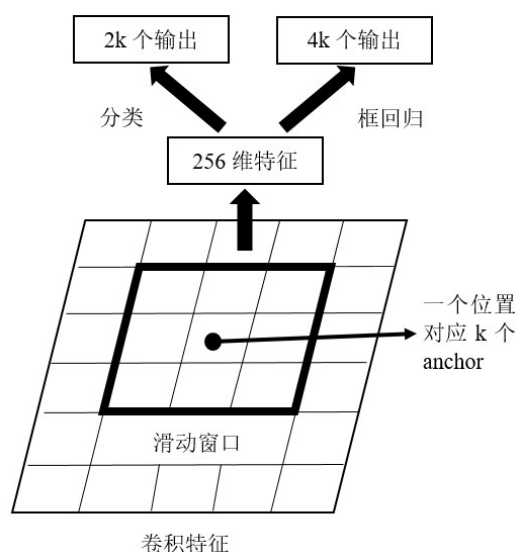
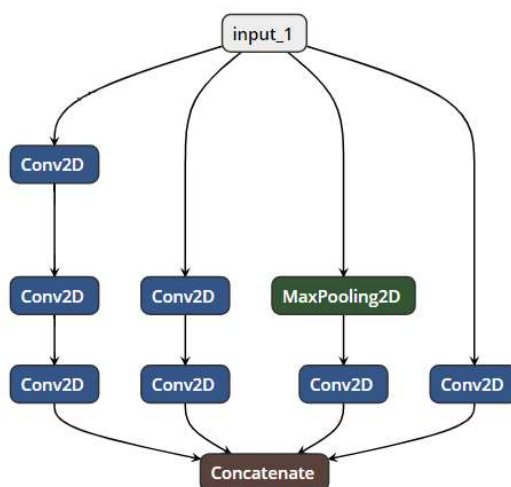


Figure 5 Inception v2 网络结构



2.2 算法设计

采用 Inception v2 网络作为算法的基本网络，其模型结构如下所示：^{[4][5]}

2.3 模型训练

在 100000 次训练后，训练损失最终降至 0.0367，模型的平均准确度达到 83.62%，当我们设定 IoU 为 0.5 和 0.75 时，平均准确度的最大值分别达到 97.44% 和 94.86%。随机抽取 100 份测试样例，模型的平均召回率达到 86.75%。

Figure 6 模型训练损失变化

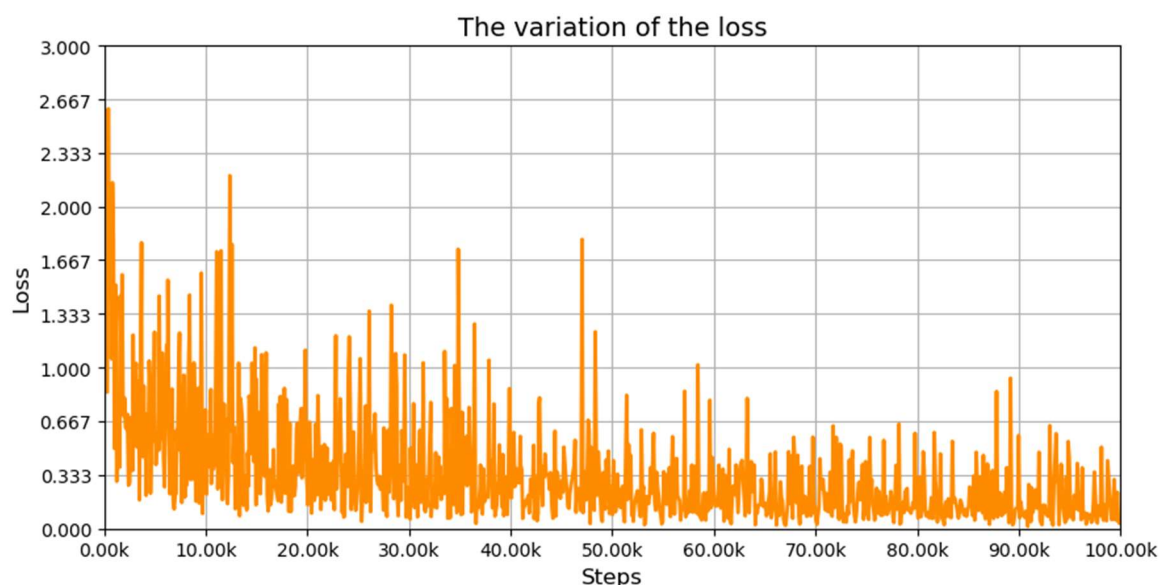


Figure 7 模型 mAP 变化

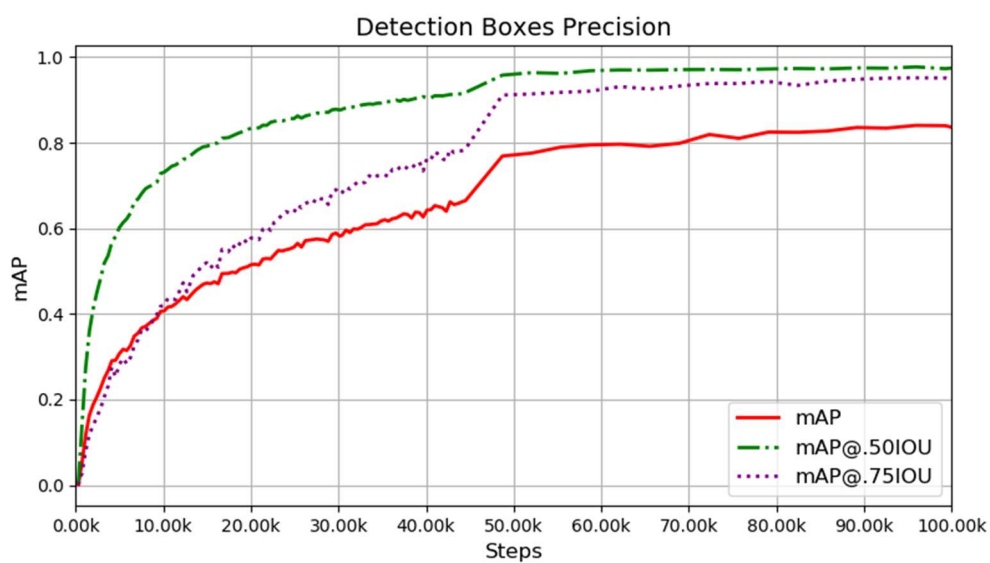
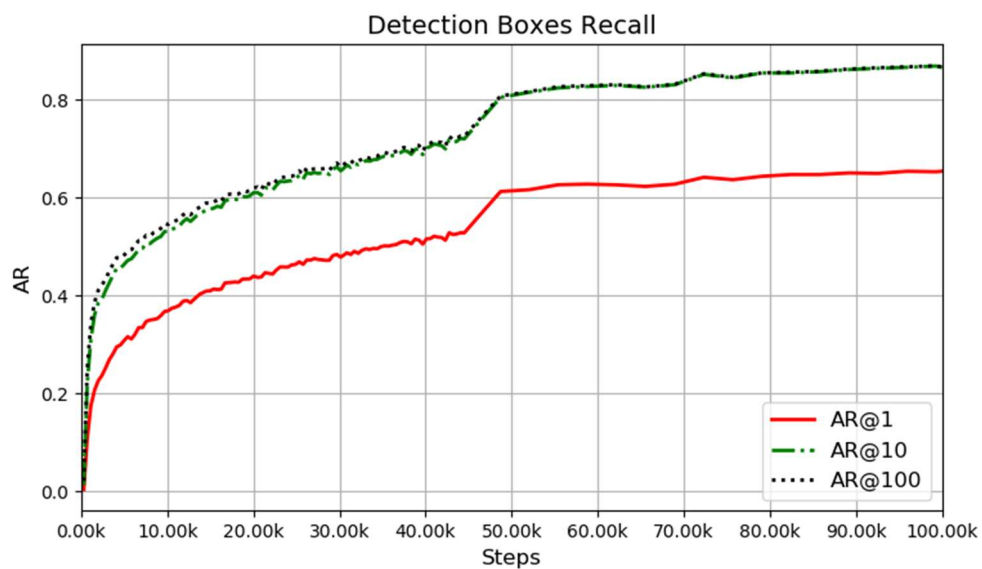


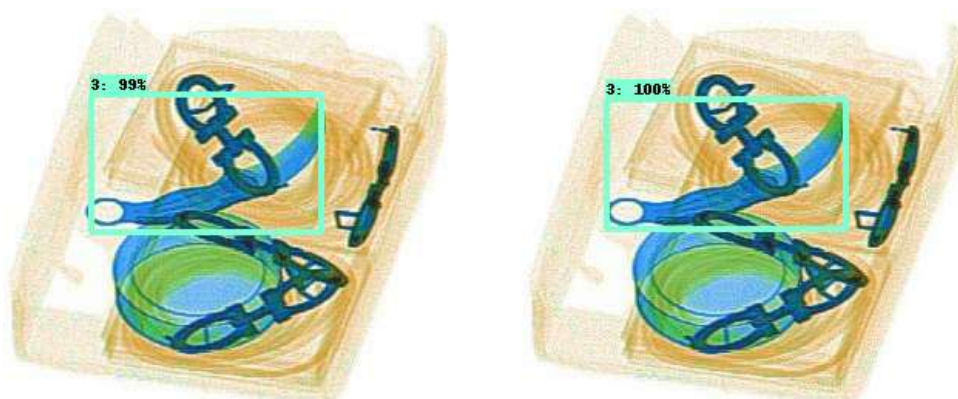
Figure 8 模型 AR 变化

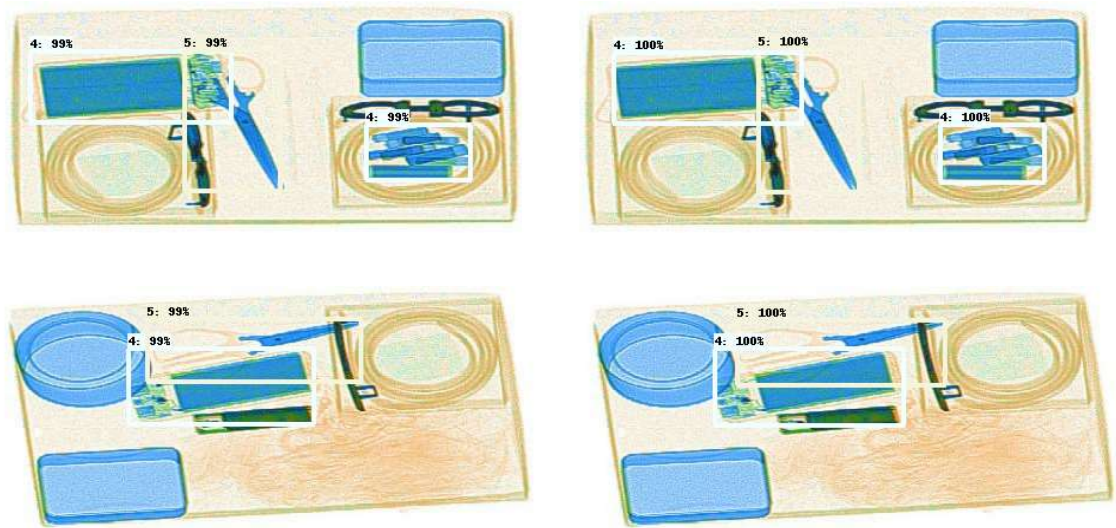


2.4 模型测试

抽取部分图像进行测试：

Figure 11 样例测试结果





从测试结果可以看到，剪刀、刀具等限制品被模型成功地标记，置信度高达 99%，在当前训练数据和训练规则下，基于当前算法的检测模型是相当成功的。

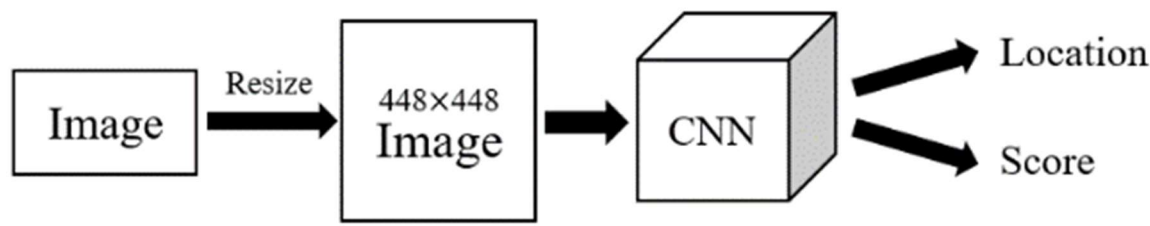
法，其不需要提前生成候选区域，而直接通过卷积网络计算目标的类别概率和坐标位置。由于所有的操作都通过一张网络进行，使得该算法具备较快的运算速度和泛化能力。

3 基于 YOLO V3 算法的检测模型

3.1 YOLO V3 算法基本原理

YOLO V3 算法属于“单阶段”算

Figure 2 YOLO V3 算法框架



YOLO V3 算法采用 Darknet-53 作为主要网络，其在 MS COCO 数据集上获得了 57.9%的平均精度。^{[6][7][8]}

Table 2 Darknet-53 网络结构

层类型		卷积核尺寸
Convolutional		(3,3,32)
Convolutional		(3,3,64)
1×	Convolutional	(1,1,32)
	Convolutional	(3,3,64)
	Residual	

	Convolutional	(3,3,128)
2×	Convolutional	(1,1,64)
	Convolutional	(3,3,128)
	Residual	
	Convolutional	(3,3,256)
8×	Convolutional	(1,1,128)
	Convolutional	(3,3,256)
	Residual	
	Convolutional	(3,3,512)
8×	Convolutional	(1,1,256)
	Convolutional	(3,3,512)
	Residual	
	Convolutional	(3,3,1024)
4×	Convolutional	(1,1,512)

	Convolutional Residual	(3,3,1024)
	Avg Pool	Global
	FC	(1000)
	Softmax	Classifier

3.2 模型训练

在训练了 50 个世代后，模型损失降至 17.7143，在测试集上的验证损失在 10 个之前不断降低，在 10 个时代之后开始增长，如下图所示：

Figure 12 模型训练损失变化
The variation of the loss

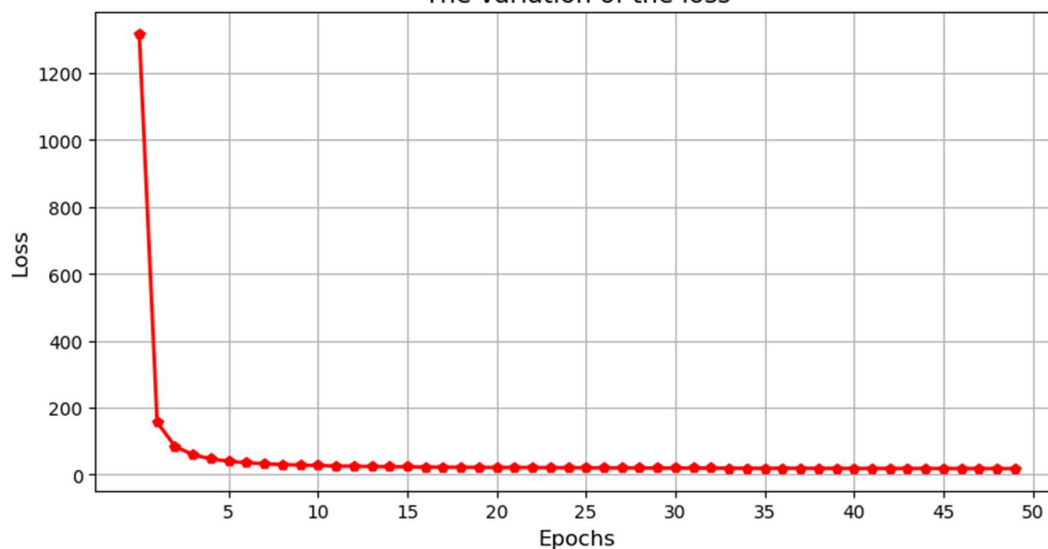
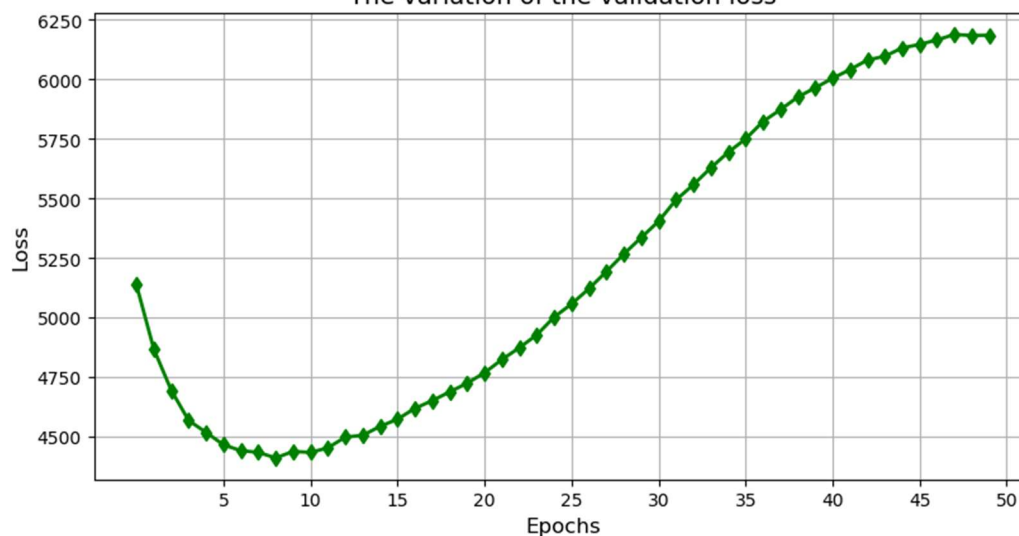


Figure 13 模型验证损失变化

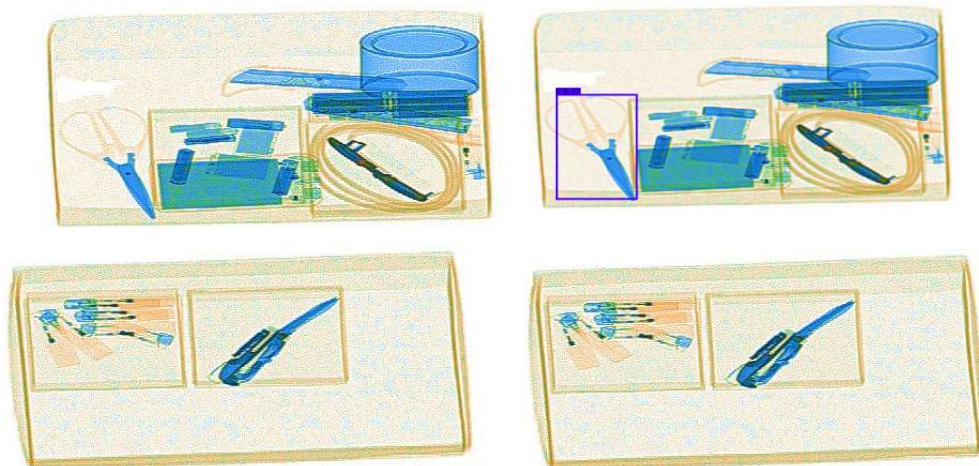
The variation of the validation loss



3.4 模型测试

取若干测试样例进行测试，测试结果如下图所示：

Figure 14 样例测试结果



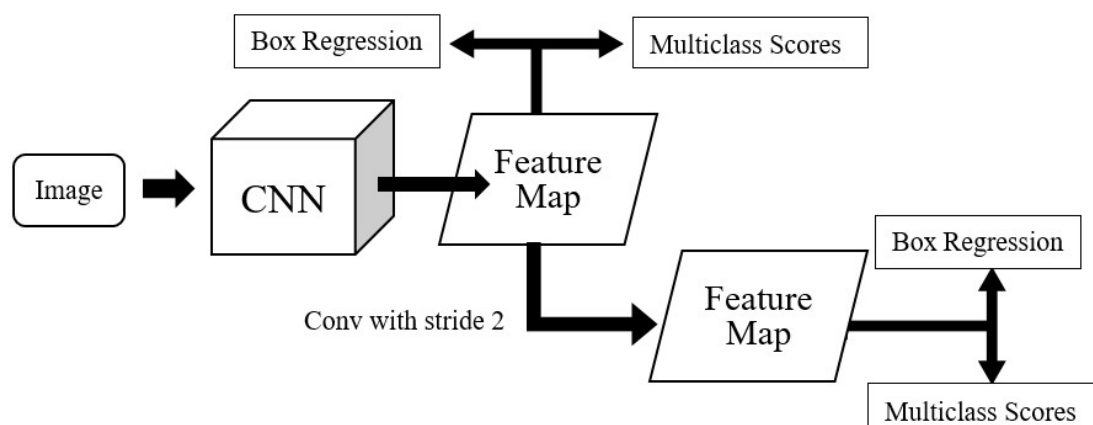
测试结果表明，模型对测试样例的识别准确率较低，仅能对部分结构简单的图像产生少量有效检测结果，甚至对于部分图像结构较为简单的样例都无法产生有效的探测框，在当前的训练数据和训练规则下，YOLO V3 算法在该检测任务上表现较差。

4 基于 SSD 算法的检测模型

4.1 SSD 算法原理

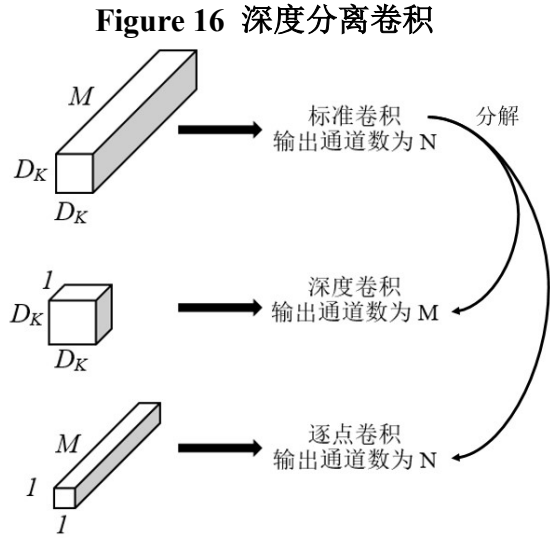
SSD 算法是一种“单阶段”检测算法，通过提取多个尺度的卷积特征进行目标检测，并且摒弃了全连接层，而直接使用卷积方式提取检测结果，同时，SSD 算法吸收了 Faster R-CNN 算法中的“anchor”理念，通过设置先验框以降低模型训练难度，提高模型准确度。^[9]

Figure 15 SSD 算法基本框架



4.2 算法设计

采用基于深度分离卷积的轻量级深度网络——MobileNet 作为算法主要网络



通过深度分离卷积，使得网络计算量减少了：

$$\frac{1}{N} + \frac{1}{D_k^2}$$

其中N为输出的通道数， D_k 为卷积核的空间维数。^[10]

MoblieNet 的网络结构如下表所示：

Table 3 MoblieNet 网络结构
(其中 Conv dw 为深度可分离卷积层)

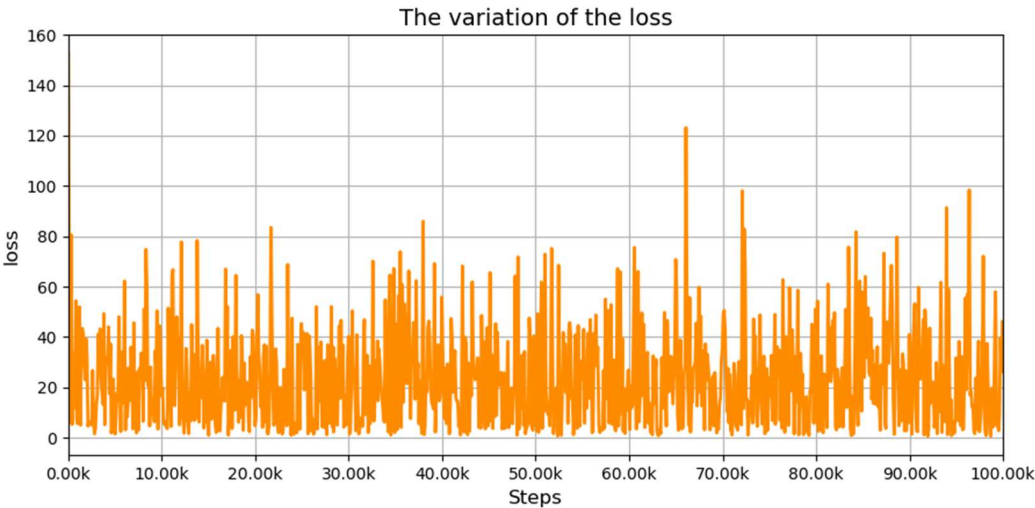
层类型	卷积核尺寸
Conv	(3,3,3,32)
Conv dw	(3,3,32)

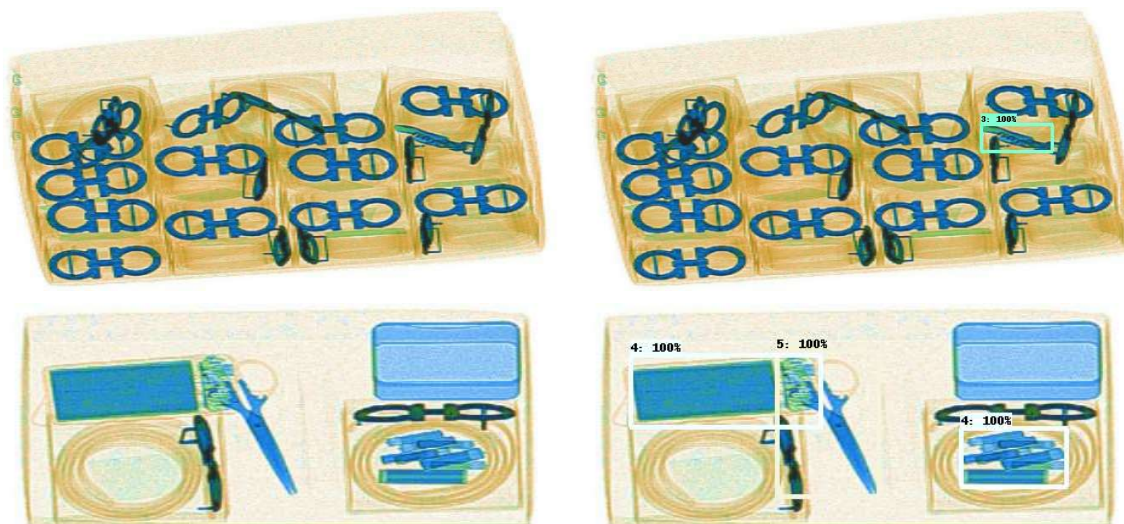
Conv	(1,1,32,64)
Conv dw	(3,3,64)
Conv	(1,1,64,128)
Conv dw	(3,3,128)
Conv	(1,1,128,128)
Conv dw	(3,3,128)
Conv	(1,1,128,256)
Conv dw	(3,3,256)
Conv	(1,1,256,256)
Conv dw	(3,3,256)
Conv	(1,1,256,512)
5×	Conv dw (3,3,512)
	Conv (1,1,512,512)
Conv dw	(3,3,512)
Conv	(1,1,512,1024)
Conv dw	(3,3,1024)
Conv	(1,1,1024,1024)
Avg Pool	Pool (7,7)
FC	(1024,1000)
Softmax	Classifier

4.3 模型训练

模型在训练过程中，训练损失呈不规则震荡变化，最终步数训练损失为46.07，模型的 mAP 和召回率始终处于较低水平，最终 mAP 和召回率为 2.17×10^{-5} 。

Figure 17 模型训练损失变化





从测试结果可以发现,在测试样例图像构造不太复杂时,模型对测试样例的识别效果较好,当测试样例的图像复杂度增加时,模型的误判率较高,在当前的训练数据和训练规则下,SSD 算法在该检测任务上的表现中规中矩。

5 综合分析及发展展望

5.1 实验结果分析

根据模型训练数据进行分析,在当前的训练数据和训练规则下,基于 Faster R-CNN 算法的模型在训练过程中,模型的 mAP 和召回率稳定增长,对于测试图像的识别具有较高的精度,基于 SSD 算法的模型在训练过程中,模型的 mAP 和召回率呈震荡变化,始终处于一个较低的水平,其对测试图像的识别精度保持在中等偏下的水平,而基于 YOLO V3 算法的模型虽然在训练过程中训练损失不断下降,但在验证集上的训练损失想减少,而后呈现爆发增长,这说明了模型出现了较

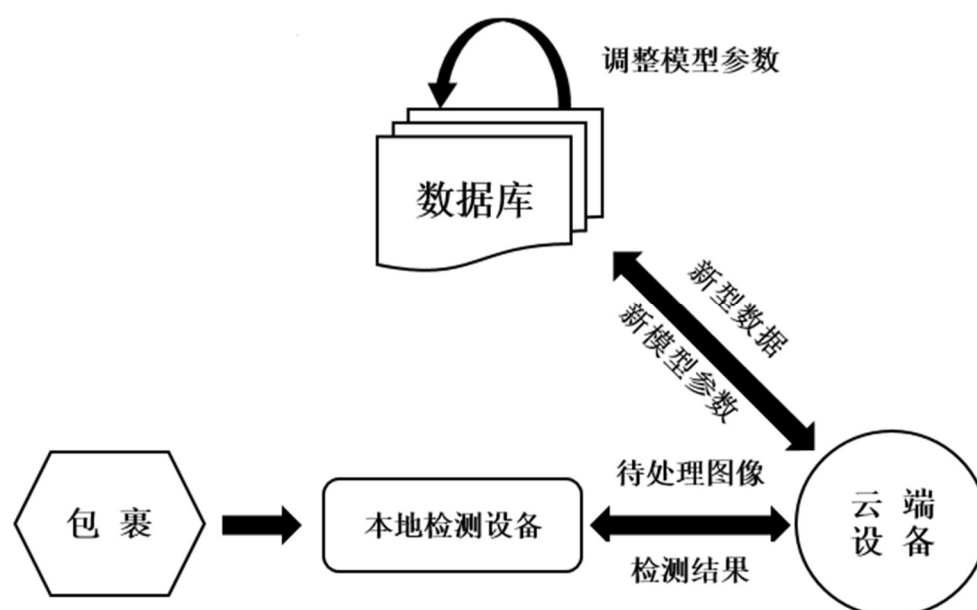
为严重的过拟合现象,其在对测试对象进行检测的过程中表现较差,因此在当前的训练数据和训练规则下, Faster R-CNN 算法为解决该检测任务的最优算法。

5.2 端云结合进行模型部署

经过前面的实验论证表明,使用自动化检测模型对安检工作进行辅助是可行的。如若使用检测模型对安检工作进行辅助,将使检测设备面临着庞大的数据处理任务,还要做到“速度快、准确率高”,这对部署端的计算能力有着极高的要求,同时,违禁品的种类和性质会随着时间推移而不断变化,这又要求对检测模型要进行及时的更新,而仅在部署端对两项要求进行贯彻,是不现实且不经济的。

云计算的高速发展和强劲表现,使得端云结合部署成为我们的良好选择,基于云计算我们可以构建以下的部署结构:

Figure 21 端云结合部署结构



本地检测设备将待测包裹数据上传至云端处理，云端高速处理数据并将结果返回，同时将已检测的数据添加进模型数据库，模型数据库定期对原始模型进行参数调整，由此不断适应新的违禁品检测任务。通过端云结合的方式，解决了本地设备算力不足，模型参数更新不够及时的问题，使整体系统呈良性循环，将不断提高系统综合检测能力。

6 结 论

本文基于三种目标检测算法对1406张包裹X光图像数据进行学习，成功建立了五种包裹携带限制品的自动化检测模型，证明了将深度学习技术应用于安检工作的可行性，同时提出了该检测算法部署的端云结合方式，

为算法的具体部署给出了建议。由于数据有限，文中所建立的检测模型仅能对五类限制品进行检测，这与现实的安检工作需求相比是不足的，为了将该检测模型真正应用到实际生活中，我们还需要做到：

- (1) 增加数据样本规模及特征复杂度，提高模型的综合检测能力；
- (2) 进一步改进模型基础网络，减小训练难度，提升网络计算速度；
- (3) 基于不同基础算法，开发多种检测模型，以配合不同的部署平台和任务需求。

伴随着目标检测技术的不断突破，应用自动化监测模型进行包裹携带品检测工作必将在不久的将来成为现实，也将极大地推动客运物流和安防行业的发展。

7 参考文献

- [1] Ren S, He K, Girshick R, et al. *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*[JJ]. 2015.
- [2] Girshick R. *Fast R-CNN*[C]// IEEE International Conference on Computer Vision. 2015.
- [3] Girshick R, Donahue J, Darrell T, et al. *Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation*[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2014.
- [4] Szegedy C, Vanhoucke V, Ioffe S, et al. *[IEEE 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Las Vegas, NV, USA (2016.6.27-2016.6.30)] 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Rethinking the Inception Architecture for Computer Vision*[JJ]. 2016:2818-2826.
- [5] Szegedy C, Liu W, Jia Y, et al. *Going deeper with convolutions*[C]// 2015.
- [6] Redmon J, Divvala S, Girshick R, et al. *You Only Look Once: Unified, Real-Time Object Detection*[JJ]. 2015.
- [7] Redmon J, Farhadi A. *YOLO9000: Better, Faster, Stronger*[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2017.
- [8] Redmon J, Farhadi A. *YOLOv3: An Incremental Improvement*[JJ]. 2018.
- [9] Liu W, Anguelov D, Erhan D, et al. *SSD: Single Shot MultiBox Detector*[C]// European Conference on Computer Vision. 2016.
- [10] Howard A G, Zhu M, Chen B, et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*[JJ]. 2017.