

# 基于优化区域卷积神经网络的机场区域检测

韩永赛<sup>1\*</sup>, 马时平<sup>2</sup>, 李帅<sup>1</sup>, 何林远<sup>2</sup>, 朱明明<sup>1</sup>

<sup>1</sup>空军工程大学研究生院, 陕西 西安, 710038;

<sup>2</sup>空军工程大学航空工程学院, 陕西 西安, 710038

**摘要** 机场区域因其特殊性对民用和军用都具有重大意义。基于机器自主识别的机场区域检测方法是目前主流的检测方法, 针对传统检测算法对机场区域遥感图像中多类别、多尺度、多视角以及复杂背景下检测鲁棒性不足的问题, 本文提出了一种优化的区域卷积神经网络检测算法。首先, 构建了一个相比传统数据集包含更多尺度、视角、类别和复杂背景等条件下的机场区域 7 类典型目标数据集并进行了优化处理, 为模型算法的监督训练和调节奠定了基础; 然后, 根据所检测目标的特性以及网络的局限性, 使用差异值法生成 anchor, 复杂负样本筛选以及加入先验判决网络对原网络进行了优化和仿真验证; 最后, 对优化的网络模型进行了测试与对比分析。实验结果表明, 本文算法在仅增加极少检测时间基础上相比原算法有更高的平均精确度, 且对各类目标的检测达到了较好的效果。

**关键词** 图像处理; 目标检测; 遥感机场区域; 区域卷积神经网络; 差异值法; 先验判决

中图分类号 TP183; TP751.1

文献标志码 A

doi: 10.3788/LOP57.101021

## Airport Area Detection Based on Optimized Regional Convolutional Neural Network

Han Yongsai<sup>1\*</sup>, Ma Shiping<sup>2</sup>, Li Shuai<sup>1</sup>, He Linyuan<sup>2</sup>, Zhu Mingming<sup>1</sup>

<sup>1</sup> Graduate School, Air Force Engineering University, Xi'an, Shaanxi 710038, China;

<sup>2</sup> Aeronautics Engineering College, Air Force Engineering University, Xi'an, Shaanxi 710038, China

**Abstract** The airport area has great significance to both civilian and military use because of its particularity. At the same time, the airport area detection method based on machine self-identification is the current mainstream detection method. Aiming at the problem of insufficient robustness of traditional detection algorithms to the detection of multiple categories, multiple scales, multiple perspectives, and complex backgrounds in airport remote sensing images, an improved regional convolutional neural network detection algorithm is proposed. Firstly, compared with the traditional data set, a typical target data set of 7 types of airport areas under more conditions such as scales, perspectives, categories, and complex backgrounds is constructed and optimized, which lays a foundation for the supervised training and adjustment of model algorithms. Then, according to the characteristics of the detected target and the limitations of the network, the difference value method is used to generate the anchor, the complex negative sample screening, and the prior decision network are added to optimize and simulate the original network. Finally, the optimized network model is tested and compared. Experimental results show that the proposed algorithm has higher average accuracy than the original algorithm on the basis of increasing only a small amount of detection time, and achieves better results for various types of targets.

**Key words** image processing; target detection; remote sensing airport area; regional convolutional neural network; difference value method; priori decision

**OCIS codes** 100.2980; 100.4996; 150.0155; 200.4260

收稿日期: 2019-11-02; 修回日期: 2019-11-29; 录用日期: 2019-12-06

基金项目: 国家自然科学基金(61701524, 61773397)

\* E-mail: 1013765061@qq.com

# 1 引言

近年来遥感成像<sup>[1]</sup>技术得到了快速发展,其成像分辨率越来越高,获取信息的能力也逐渐增强,因此如何对遥感图像进行充分挖掘利用成为了目前计算机视觉和遥感研究的一个热点话题<sup>[2]</sup>。机场作为军用和民用的重点建设目标,目前已经成为了目标检测领域的“宠儿”<sup>[3]</sup>。

许多研究人员对这一问题进行了相关研究并在文献中发表了有价值的研究结果。Zhao等<sup>[4]</sup>结合Hough变换和基于图形的视觉显著性(GBVS)方法结合进行机场检测,Tao等<sup>[5]</sup>结合了分割区域和尺度不变特征变换(SIFT)特征结合统计方法进行机场检测,大大加速了机场目标的检测速度但同时都存在人工设计特征的局限性。Zhu等<sup>[6]</sup>从几何显著性和局部熵结合的显著图上得到感兴趣区域(ROIs),再利用AlexNet识别网络迁移学习来进行机场识别,Zhang等<sup>[7]</sup>在先验知识的基础上提出了一种机场呈线性特征的区域建议方法,并结合卷积神经网络(CNN)来进行机场的检测,发挥了CNN强大的分类能力,在一定程度上提高了检测率,但是对于区域建议(region proposal)的选取依然都存在传统方法的局限性。Chen等<sup>[8]</sup>将深度置信网络(DBN)用于飞机检测,发挥了DBN很强的分类能力,但检测性能较低且时间成本较高。戴陈卡等<sup>[9]</sup>将Faster R-CNN和多部件结合方法用于机场飞机检测,大大提高了检测性能且降低了时间成本,但由于数据集等条件限制,对小目标飞机特征提取不足,检测不理想。朱明明等<sup>[10]</sup>在卷积神经网络的基础上应用特征融合和软判决的方式对其进行了改进,有效地提高了对小目标的检测率和遮挡问题。单倩文等<sup>[11]</sup>基于改进的多尺度特征图算法对网络提取的特征图进行了融合、降维以及与注意力机制的结合,较为有效地提高了检测精度。张超等<sup>[12]</sup>基于困难样本挖掘与残差网络结合的方式进行目标检测,在一定程度上提高了检测精度且未增加较大检测时间。但是众实验多是建立在自然图像基础上或相似拍摄距离和固定的垂直视角下的单目标或者目标种类、数量较少且背景较为简单的遥感图像基础上,对不同拍摄距离、不同视角和较为复杂背景下的目标数据集较少,而不同高度,不同视角等得到的同一目标图像差别很大<sup>[13-14]</sup>,这也是造成遥感图像检测鲁棒性不足的原因。

针对机场、民航飞机、战斗机、运输机、直升机、

油罐以及桥梁的不同高度,尤其是超高空下的机场快速搜索检测和远距离下的其余目标精确检测,以及多视角下的检测问题,本文基于目前应用广泛的深度学习技术,提出了一种通过自主构建包含更多尺度、视角的目标数据库和优化区域卷积神经网络进行自学习地特征提取并训练出相应权重模型来对不同距离、不同视角、复杂背景下目标进行精确的分类和定位的方法。最后通过实验以及与其他算法和原网络的对比,证明了本文算法的有效性和有意义性。

## 2 区域卷积神经网络的特点与优化

### 2.1 区域卷积神经网络的特点及其局限性

#### 2.1.1 网络的构成特点以及实现机场区域检测的原理

区域卷积神经网络是目标检测中使用最多的神经网络,其中Faster R-CNN<sup>[15]</sup>是该类网络的集大成者,其检测步骤主要分为两步:区域建议和检测。其检测步骤的主要流程如图1所示。

1) 网络引入大容量卷积神经网络提取深层特征

现在的目标检测任务中无论是目标的本身复杂性,如旋转、尺度变换、视角变换,还是背景的复杂性,如相似物干扰、对比度变换、噪声干扰等,传统的人工设计特征以及一般的特征提取方法想要获得鲁棒性强的特征困难重重。而大容量卷积神经网络具有强大的特征提取能力,它依靠机器大量运算而获得更加鲁棒的特征,避免了由人工设计特征的经验局限性。其主要包括卷积层、激活函数层以及池化层。

卷积层对输入图片进行卷积操作。相关参数为 $S_{\text{kernelsize}}$ 代表卷积核尺寸; $L_{\text{pad}}$ 代表扩边处理的长度(如 $L_{\text{pad}}=1$ 即在矩阵周边填充了一圈0); $S_{\text{stride}}$ 代表步长(即每次卷积核移动的长度)。卷积尺寸变换公式为

$$S_{\text{outputsize}} = \frac{S_{\text{inputsize}} - S_{\text{kernelsize}} + 2 \times L_{\text{pad}}}{S_{\text{stride}}} + 1, \quad (1)$$

式中: $S_{\text{outputsize}}$ 代表输出图像尺寸; $S_{\text{inputsize}}$ 代表输入图像尺寸。

激活函数层主要作用为增加神经网络模型的非线性,让神经网络不仅局限于矩阵的线性运算。

池化层主要对经过卷积操作的矩阵进行降维处理。

2) 区域建议网络(RPN)用于更快地进行区域建议

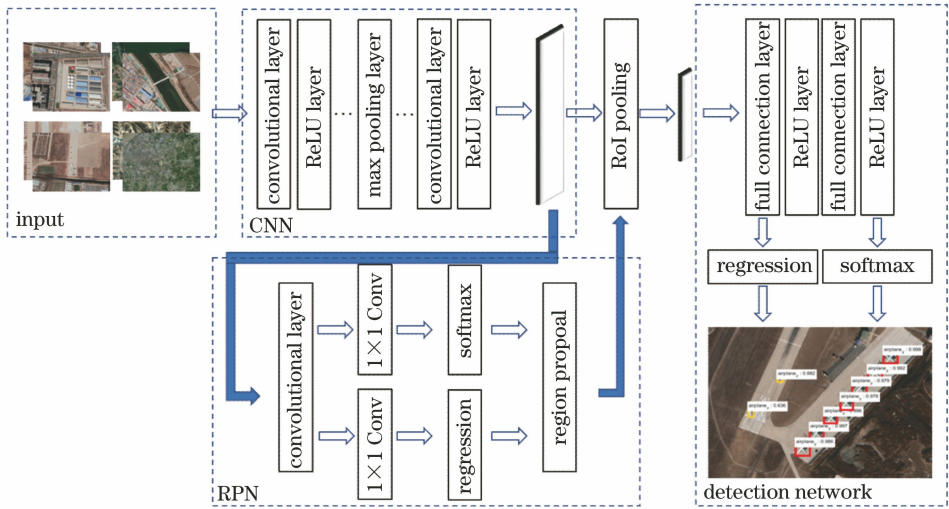


图1 检测步骤的主要流程框架图

Fig. 1 Main process framework of detection

区域建议网络使得建议区域的提取过程放在特征图上,显而易见,更低分辨率的特征图代表了更小的计算量,因此大大降低了由网络在原图上滑动窗口以及建议框重叠部分被多次重复提取特征所带来的时间和存储空间上的消耗,在很大程度上改善了模型设计的检测速度性能。

区域建议网络主要用于建议区域的生成,首先会依据不同的比例以及面积对每个锚点来生成9个锚框(anchor),在原算法中取值为 $128^2$ 、 $256^2$ 、 $512^2$ 像素点三个面积以及1:1、1:2、2:1三个长宽比,共9个尺度。然后通过一定的规则对其进行过滤,只剩下符合要求的锚框,再通过分类函数(softmax)判断是目标还是背景,即只进行是或不是目标的二分类判断;同时定位回归函数用来修正锚框的位置坐标,进而较为精确地建议区域。

其中区域建议网络的权重更新以达到能够预测框的能力是通过对比真实目标标注框(ground truth)与锚框之间的偏移量来学习的。相关公式为

$$\begin{cases} \Delta x = (x^* - x_a) / w_a, \Delta y = (y^* - y_a) / h_a \\ \Delta w = \log(w^* / w_a), \Delta h = \log(h^* / h_a) \end{cases}, \quad (2)$$

式中: $x^*$ 、 $y^*$ 、 $w^*$ 、 $h^*$ 分别为真实目标标注框的中心点位置坐标、宽和高; $x_a$ 、 $y_a$ 、 $w_a$ 、 $h_a$ 分别为锚框的中心点位置坐标、宽和高。

3) 感兴趣池化(RoI Pooling)层以及全连接层构成最终的检测网络

感兴趣池化层主要对区域建议网络生成的建议区域和由模型构建的卷积网络上得来的特征图生成

区域特征图,然后进行尺度统一化处理,即假设最后输入全连接层的尺寸要求为 $h \times w$ ,则池化层会把特征图分为 $h \times w$ 个网格,对每个网格元素进行取最大值处理,得到统一的 $h \times w$ 的特征图尺寸。

全连接层主要进行目标的具体分类以及目标区域位置的进一步精确定位,最终得到需要的输出信息。RoI Pooling层和全连接层在算法框架中的位置如图1所示。

#### 4) 网络实现机场区域检测的原理

本文所检测的机场区域目标均以光学遥感图像为检测对象的源,以上述区域卷积神经网络为检测手段,以最终输出带有定位框和置信度的检测图像为输出。当输入一张待检测的图片后,网络首先将会对图片进行批处理成固定尺寸(网络支持输入任意大小的图片,此项操作是由于全连接层对尺寸有要求),然后进入特征提取网络进行自主特征提取得到一组特征图(VGG16中每组特征图数量为512)。此后,特征图将分为两步:①进入区域建议网络进行建议区域提取;②进入感兴趣池化层与区域建议网络提取的建议区域一起形成池化后的特征图。最后,对感兴趣池化后的特征图进行全连接操作,利用分类损失和定位损失函数完成目标的精确分类和定位并输出最终的检测结果图,完成对机场区域目标的检测。图2为步骤示意图。

#### 2.1.2 网络的局限性

合理的区域建议是取得良好目标分类以及定位的基础之一。从上述网络特点可以看出,锚框是进行区域建议的关键,而传统的锚框为 $128^2$ 、 $256^2$ 、

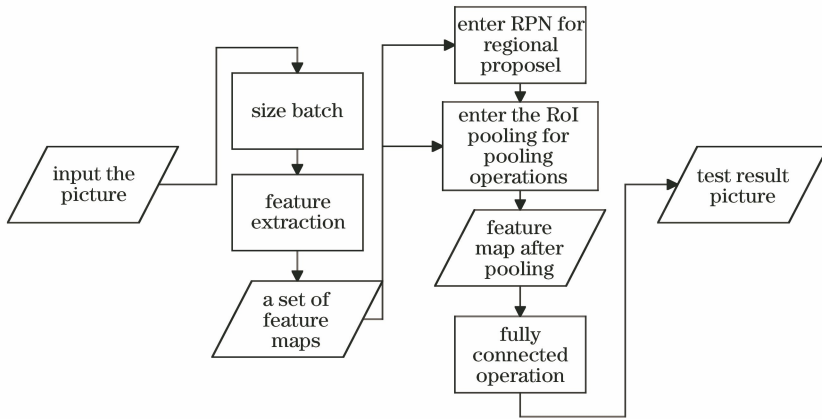


图 2 机场区域检测步骤示意图

Fig. 2 Schematic of airport area detection

512<sup>2</sup> 像素点三个面积以及 1:1、1:2、2:1 三个长宽比组合成的 9 个固定尺度, 无法很好地适应多类、多尺度的目标。

网络在训练过程中容易出现正负样本失衡问题, 尤其是对复杂背景下的目标检测。在复杂背景目标检测任务中, 背景的数量是目标数量的几倍甚至几十倍, 若不进行良好的均衡控制会导致训练对负样本(背景)敏感而对正样本(目标)不敏感, 从而大大降低检测性能甚至网络失效。

传统的遥感目标检测通常为单类目标的检测, 针对多类遥感目标检测(即一张图像上可能具有多类目标)时, 网络缺乏类间共存合理性判断, 当区域卷积神经网络在进行多类目标检测时, 检测的重点放在一个特征分属于各个目标种类的概率大小, 所以可能存在不符合先验经验的误检出现, 比如在一张图上检测到飞机的同时检测到

机场, 这是不可能的, 因为当检测到飞机时, 卫星拍摄高度会降的很低, 这时是不可能图像中出现一个完整机场的。

## 2.2 网络的改进优化与仿真

### 1) 基于差异值的锚框生成规则的改进

在锚框的生成规则的改进中, 传统且最普遍的方法为人工制定锚框的面积与比例, 简便易行且通常可以达到较好的效果, 但是由于本文所构建的数据集较复杂, 目标尺度众多, 甚至同类目标可能存在几十种尺度, 因此在考虑算法复杂度的基础上人工制定出几种尺度来符合大部分目标的尺度是很难的。若制定的尺度太小则提取不到目标的全局特征, 不利于分类; 若制定的尺度太大则不利于获得目标的精确位置信息且会增加卷积等计算量, 拖慢检测效率。在此, 本文使用差异值算法得到更为精确的锚框尺度。其步骤如表 1 所示。

表 1 差异值算法生成锚框步骤示意

Table 1 Schematic of difference value algorithm generates anchor step

Difference value algorithm generates anchor box

- Step 1: Extract the area and proportion information of the ground truth of some targets in each type of target from the regional proposal network as a sample.
- Step 2: The information extracted from various targets is transformed into a two-dimensional European space.
- Step 3: Initialize 9 anchor boxes randomly (the number selection is modeled after the Faster R-CNN detection algorithm. Too much is easy to multiply the calculation amount, and too few is not easy to represent the full scale of the target) and compare the 9 anchor boxes with all of the selected samples ground truth information and calculate the difference value of each box.
- Step 4: The ground truth with small difference value is divided into a combination around the corresponding anchor box.
- Step 5: Calculate the average size of the ground truth in each combination as a new anchor box.
- Step 6: Repeat the above steps until the difference does not change much after each iteration, and get the best 9 anchor boxes.

在进行差异值计算函数设计时, 考虑以机器生成的锚框与真实目标标注框的重合度作为自变量,

重合度越大其差异值应该越小。定义差异值的取值范围是(0, ∞), 当差异值为 0 时, 代表锚框与真实目

标注框完全重合,差异值越大代表锚框与真实目标标注框重合度越小。此外当重合度从 0 递增时,差异值开始快速衰减以对其敏感而后衰减速度越来越慢,当重合度接近 1 时有变化不大的差异值。综合考虑,差异值函数定义为

$$D = -\ln[C(w_{gt}^i, h_{gt}^i, w_a^k, h_a^k)], \quad (3)$$

$$C = U_{IoU(a,gt)} = \frac{S(A \cap G)}{S(A \cup G)}, \quad (4)$$

式中: $D$  为差异值; $C$  为重合度; $w_{gt}^i$  为第  $i$  个真实目

标注框的宽; $h_{gt}^i$  为第  $i$  个真实目标标注框的高; $w_a^k$  为第  $k$  个锚框的宽; $h_a^k$  为第  $k$  个 anchor 框的高; $U_{IoU(a,gt)}$  为锚框与真实目标标注框的交并比; $S(A \cap G)$  为锚框与真实目标标注框相交的面积; $S(A \cup G)$  为锚框与真实目标标注框相并的面积。

以机场目标为例,差异值算法生成的锚框以及传统锚框的对比如图 3 所示,与目标贴合度最高的为基于差异值算法生成的锚框,其余框为传统的锚框。

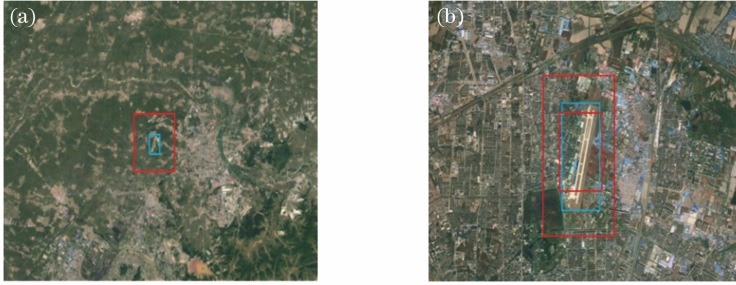


图 3 差异值算法生成的锚框以及传统锚框

Fig. 3 Anchor generated based on the difference value algorithm and traditional anchors

仿真结果如表 2 所示,其中 T1 代表本文提出的差异值算法,mAP 表示多类别平均检测精度。

表 2 T1 改进算法与原算法性能对比表

Table 2 Comparison of T1 improved algorithm and original algorithm performance

Method	mAP / %	Mean time / s
Faster R-CNN	67.5	0.142
Faster R-CNN+T1	70.3	0.142

通过图 3 对比可以发现,本文算法生成的锚框更加适合目标的尺度,生成的尺度已经不局限于原算法中的 9 种固定尺度,而是能更加贴合目标的 9 种尺度,进而在一定程度上改善了 Faster R-CNN 中固定的锚框尺度不能很好适应多类、多尺度目标以及人工设计锚框尺度范围不易较好地适用于多类别、多尺度目标的局限性;且通过抽取部分数据集进行仿真验证了差异值算法改进的有效性,通过改进算法网络取得了更好的多类平均检测精度,同时因为改进算法主要应用于训练过程中锚框尺度的生成,因此未造成检测时间的增加,保证了检测效率。

### 2) 筛选复杂负样本以改善训练样本平衡问题

传统的定义正样本的方法为:在一组框中得分最高的框以及与真实目标标注框的交并比(IoU)大于 0.7 的框。定义负样本的方法为:在一组框中与真实目标标注框的交并比小于 0.3 的框。而其余样本框对于训练是没有贡献的。本文所研究的对象具有多尺度特性且多数为较小尺度,因此负样本会很

多,很容易引起训练中的正负样本比例失衡问题,进而导致训练对负样本(背景)敏感而对正样本(目标)不敏感。在此,建立筛选网络用于从区域建议网络的非极大值抑制(NMS)算法计算的交并比中选取值在(0,0.15)间的复杂负样本进行训练,以减少可选负样本的量来改善正负样本失衡问题,并通过仿真实验证明了复杂负样本选取的合理性。

图 4 为准确率-召回率(P-R)仿真曲线对比,其中,横坐标代表精确率,表示所有被检测为正样本中是正确预测的比例;纵坐标代表召回率,表示被正确识别出的正样本数占全部真正样本数的比例。

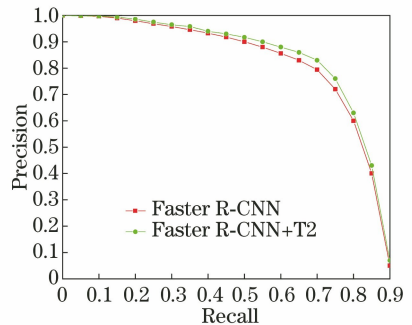


图 4 加入 T2 的改进算法与原算法 P-R 曲线对比图

Fig. 4 Comparison of the improved algorithm of adding T2 and the original algorithm P-R curve

仿真结果如表 3 所示,其中 T2 代表本文提出的筛选网络。

表3 加入 T2 的改进算法与原算法性能对比表

Table 3 Comparison of improved algorithm and original algorithm performance of adding T2

Method	mAP / %	Mean time / s
Faster R-CNN	67.5	0.142
Faster R-CNN+T2	68.8	0.143

通过 P-R 曲线以及结果对比表可知,筛选网络的加入使得可选负样本控制在交并比值在(0,0.15)之间的负样本而正样本的选取不进行约束,从而在训练过程中正负样本比例得到了一定改善,并且网络可以从复杂负样本中学到更多的区分性特征信息。通过抽取部分数据集进行仿真发现,网络拥有更好的检测曲线特征,相比原算法在多类平均检测精度方面有一定提升,同时筛选网络应用于区域建议网络之中,贯穿于训练和检测的全过程,因此对检测时间会有一定影响。但是计算机读取交并比的值以及进行排序筛选的速度是很快的,因此从仿真结果来看,筛选网络的加入仅增加了非常少的检测时间,可以证明筛选网络以非常小的代价换取了较大的性能提升。

3) 加入先验判决网络来进行类间共存合理性判断

表4 先验判决算法步骤

Table 4 Prior judgment algorithm steps

## A priori decision implementation steps

- Step 1: Read the classification results of the detection network from the log file (where the labels are assigned to the values 0, 1, 2, ..., 6 in the order in Table 6) and the corresponding confidence levels.
- Step 2: If multiple types of labels are detected and the product of the label values is 0, then Step 3 is performed, otherwise the label name is directly output.
- Step 3: Compare the average of the detection confidence of the target with a non-zero label to the average of the target detection confidence with a label value of 0 to obtain a label with a larger average confidence value. If the target average confidence level with a label value of 0 is large, 0 is output, otherwise all other non-zero label values are output.
- Step 4: Read the label value in Step 3 and output the corresponding label name.

算法的核心思想为消除在较低的卫星视角进行飞机类目标、桥梁以及油罐目标的检测时可能出现的将其他背景检测为机场的误检,此外也消除在进行机场目标检测时可能出现的将其他背景检测为飞机类目标、桥梁以及油罐目标的误检。并以 T3 代表本文先验判决算法。

通过抽取部分数据集进行仿真得到平均检测时间如下:Faster R-CNN 为 0.142 s;Faster R-CNN+T3 为 0.143 s。

工作特性曲线(ROC)如图 6 所示。其中,横坐标代表假正率(虚警率),表示所有负样本中错误预测为正样本的概率;纵坐标代表真正率(命中率),表示所有

在对 Faster R-CNN 进行训练测试时,得到了一些不符合常规的检测结果如图 5 所示(在每个目标上加入了置信度的数值),误将类似机场几何特性的区域错误地检测为机场,但是这在现实中是不可能存在的。因此有必要加入先验判决来规避类似的不符合常规的误检。



图5 误检示意图以及局部放大

Fig. 5 Schematic diagram of false detections and partial magnification

通过论证分析可以发现,出现此类误检时通常误检目标置信度较低(一般低于0.7)。因此,本文在检测网络给出类别信息以及置信度之后,在输出最后检测结果之前加入先验判决用来消除这种误检。具体步骤如表 4 所示。

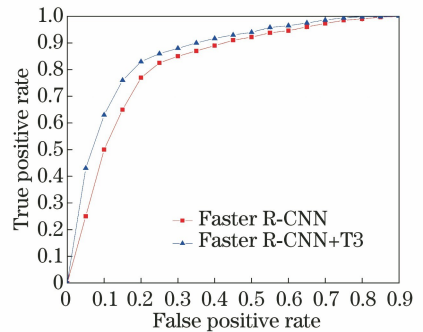


图6 T3 改进算法与原算法的 ROC 曲线对比图

Fig. 6 Comparison of ROC curve between T3 improved algorithm and original algorithm

正样本正确预测为正样本的概率。由 ROC 仿真曲线的对比可知,加入先验判决后,对于相同真正率来说,网络有更小的假正率,究其原因该算法通过在最后输出结果阶段增加类间共存合理性判断消除了不符合先验情况的检测,而并没有影响原算法应有的检测能力,因此在具有相同命中率(检测能力一致)时,网络拥有了更低的虚警率,证明了先验判决改进的有效性。同时,该方法在进行计算时,计算机提取日志文件、运

行算术运算的速度很快,因此仅增加了很少的检测时间,在获得更好检测率的同时保证了检测效率。

### 3 基于优化算法的机场区域目标检测

#### 3.1 数据集的构建及优化

##### 3.1.1 数据集的构建

本文实验用到的数据集为 VOC2007<sup>[16]</sup> 格式,具体流程如图 7 所示。

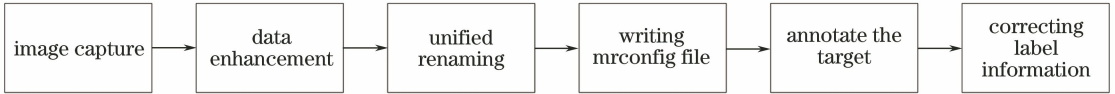


图 7 数据集构建流程图

Fig. 7 Flow chart of data set construction

本文主要面向的目标检测图像为遥感图像,而且已有文献说明 Google map 上获取的遥感图像应用于其他传感器时依然有良好的泛化能力<sup>[17]</sup>。因此应用 Google Earth Pro 软件提供的遥感图像作为图像截取搜集工作的源,在其中截取不同高度、不同视角以及不同背景下的目标遥感图片。使得神经网络能自主学习多尺度、多视角以及不同背景下的目标。通过这种方式对 200 多个机场

以及场内的飞机、油罐和众多桥梁进行了图像的截取,得到 4235 张原始图像。所得图像集的目标拥有多尺度、多视角、多类别以及简单/复杂背景特性,图像集的完备性较好。其中,共有机场图像 1081 张、民航飞机图像 1130 张、战斗机图像 363 张、直升机图像 413 张、运输机图像 402 张、桥梁图像 380 张以及油罐图像 466 张。与传统数据集的特色对比如表 5 所示。

表 5 实验数据集与传统数据集特色对比

Table 5 Comparison of experimental data sets and traditional data sets

Item	Traditional remote sensing target detection data set	Experimental target detection data set
Category	Single class	Multi-class
Scale	Medium/large	Small/medium/large scale (especially focusing on small scale targets)
Perspective	Vertical viewing angle	30°, 60°, 90°, etc. Multi-viewing angle
Background	Simple background	Focus on target detection in complex backgrounds (especially airport backgrounds)

部分数据集示意如图 8 所示。

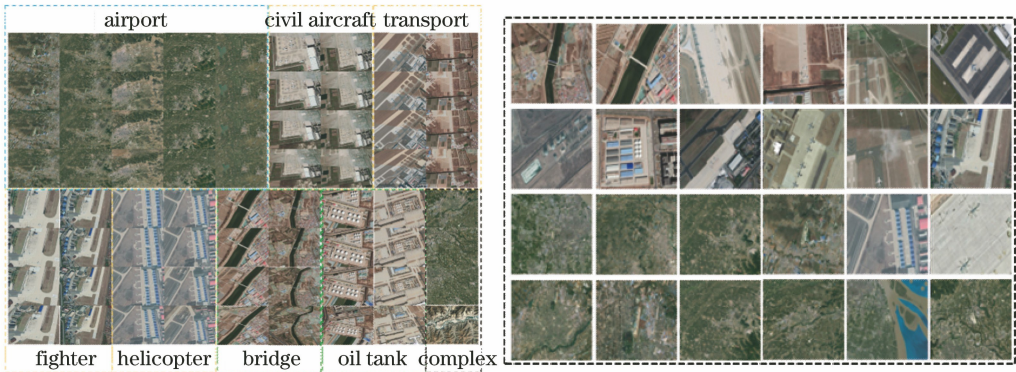


图 8 部分原始数据集

Fig. 8 Partial raw data set

针对各类目标制定的标签如表 6 所示。

表 6 标签及其对应目标对照表

Table 6 Label and its corresponding target comparison table

Label	airport	airplane_mh	airplane_z	airplane_zs	airplane_y	bridge	oiltank
Object	Airport	Civil aircraft	Fighter	Helicopter	Transport	Bridge	Oil tank

3.1.2 对数据集的优化处理

1) 数据增强处理

为了使神经网络学习到更多的目标信息,同时也为了减小因数据集容量较小而可能带来的过拟合问题,本文应用翻转、旋转、亮度变换以及增加噪声的常规方式进行了数据增强,使得网络能够学习更多的目标信息,增强目标检测的鲁棒性。考虑到硬件性能的限制,数据集不能过大,因此采用随机扩充的方法使得每张图片随机地被某种方法处理,图像数据量扩充一倍。

2) 对标注信息的快速纠错处理

使用的标签数量为 7,且涉及未进行数据增强前多达几千幅的原始遥感图片,因此在进行人工标注时,由于误操作、疲劳等因素会不可避免地会出现错误标记。错误标记包括漏标以及误标,漏标需要人工反复核对才能消除,但是误标是可以规划更快的定位方法的。误标的原因是在进行标注,尤其是小目标的标注时,大量注意力放在更好地调整矩形框尽量准确与其相切,且大量标注软件的标签信息是显示在目标标注框内的,当标注框本身较小时,就容易造成标签的不可完全可视化。

在此本文提出如下方式进行误标信息的快速纠正:首先,每类目标的图片放在单独的文件夹中,对一张图上包含多类目标时,按表 6 标签顺序放在排序最前的相应目标文件夹,对于标注信息.xml 文件放在对应图片所在文件夹;然后,对每个

目标所在文件夹批量搜索标注的.xml文件中是否包含其余 6 个目标的信息,若存在则进行检查修正工作。经过优化处理,花费不到 10 min 的时间搜索并更正了 100 多张标注错误的图像,节省了大量对可能存在误检标签的文件检查时间且将修正率提高到 100%,为下一步取得更好的训练效果做出了贡献。

3.2 仿真环境以及参数设置

1) 平台环境:处理器为 Intel(R) Core(TM) i7-7700 CPU@3.60 GHz 3.60 GHz,安装内存为 16.0 GB,操作系统为 Windows 64 位,显卡为 NVIDIA GeForce GTX 1060 6 GB,实验框架为 Caffe,Matlab。

2) 参数设置:采用深度学习的 Caffe 框架,全连接层和卷积层权重选取 VGG16<sup>[18]</sup> 预训练网络初始化,其余新层采用均值为 0、标准差为 0.01 的高斯分布随机初始化;初始学习率设置为 0.001,动量设置为 0.9,权重衰减设置为 0.0005;阈值设置为 0.7;训练次数为 40000。

3.3 模型的训练及结果分析

3.3.1 模型的训练

共享卷积层可以避免多卷积层重复训练而产生大量计算冗余进而大大提高训练和检测效率,在此采用交替优化策略训练网络实现卷积层共享。训练步骤如表 7 所示,其中样本数之比为训练集:验证集:测试集=7:2:1。

表 7 算法步骤

Table 7 Algorithm steps

Algorithm steps
Step 1: Train the region proposal network separately, initialize the weights by the pre-trained model, and adjust the parameters in an end-to-end manner to give a proposal region.
Step 2: Train the detection network separately. The region area for training comes from Step1. The weights are initialized using a pre-trained model.
Step 3: Use the parameters of the Step2 detection model to initialize the regional proposal network while fixing the convolutional layer, and adjust only the regional proposal network parameters.
Step 4: Use the proposal area output from Step3 as the input to the detection network, while keeping the shared convolutional layer fixed and fine-tune the remaining detection network parameters.

3.3.2 本文算法的整体检测结果与分析

通过训练得到的测试结果如表 8 所示。整体上看,经过训练模型对机场以及民航飞机的检测效果

较好,平均检测精度 (AP) 分别为 80.8415% 和 84.8188%;且具有较高的检测效率,其中仿真得到平均检测时间为 0.145 s。此外,对战斗机、运输机、



表 8 各目标测试结果汇总表

Table 8 Summary of each target test results

Object	Airport	Civil aircraft	Helicopter	Fighter	Transport	Oil tank	Bridge
AP / %	80.8415	84.8188	70.0974	62.1441	71.0077	73.5869	68.7273

直升机、油罐以及桥梁(主要指机场附近)的检测有待加强,这是由于一些保密措施、地理因素等,战斗机、运输机、直升机、油罐以及桥梁的数据相比机场和民航飞机的遥感图像要少很多,而用于训练的基数少,最终算法可以提取到的区分性特征鲁棒性势

必会下降;此外对于战斗机来说有比其他目标更小的尺度,在本文的高距离拍摄视角下所对应的像素点更少,对检测造成的困难更大。

可视化各类目标部分检测结果示意与分析如表 9~11 所示。

表 9 机场的检测结果

Table 9 Airport test results

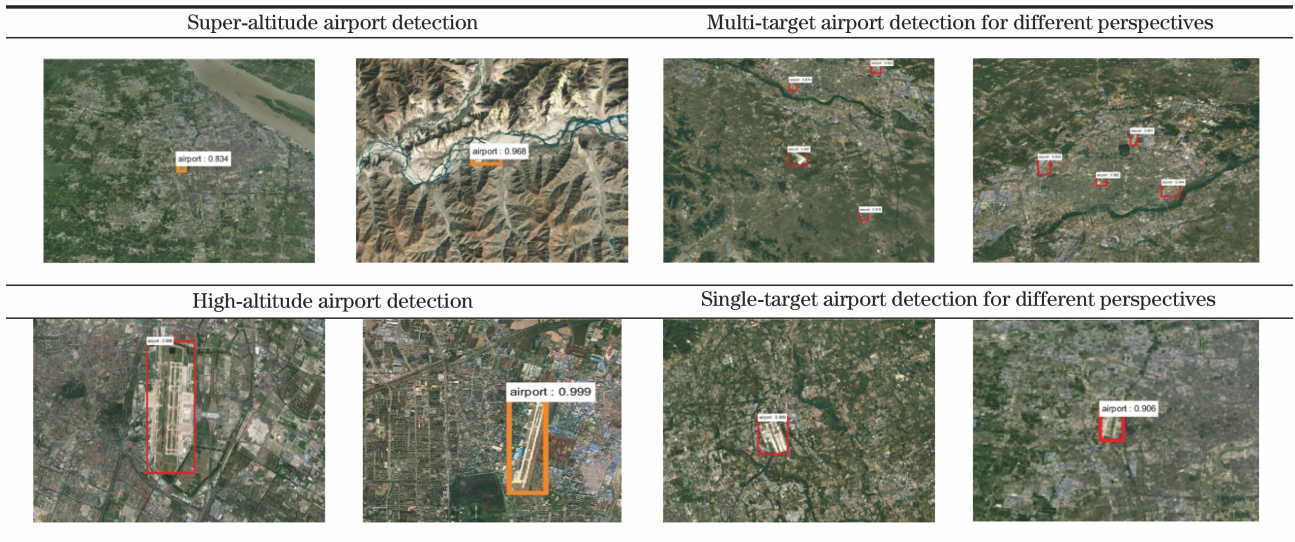
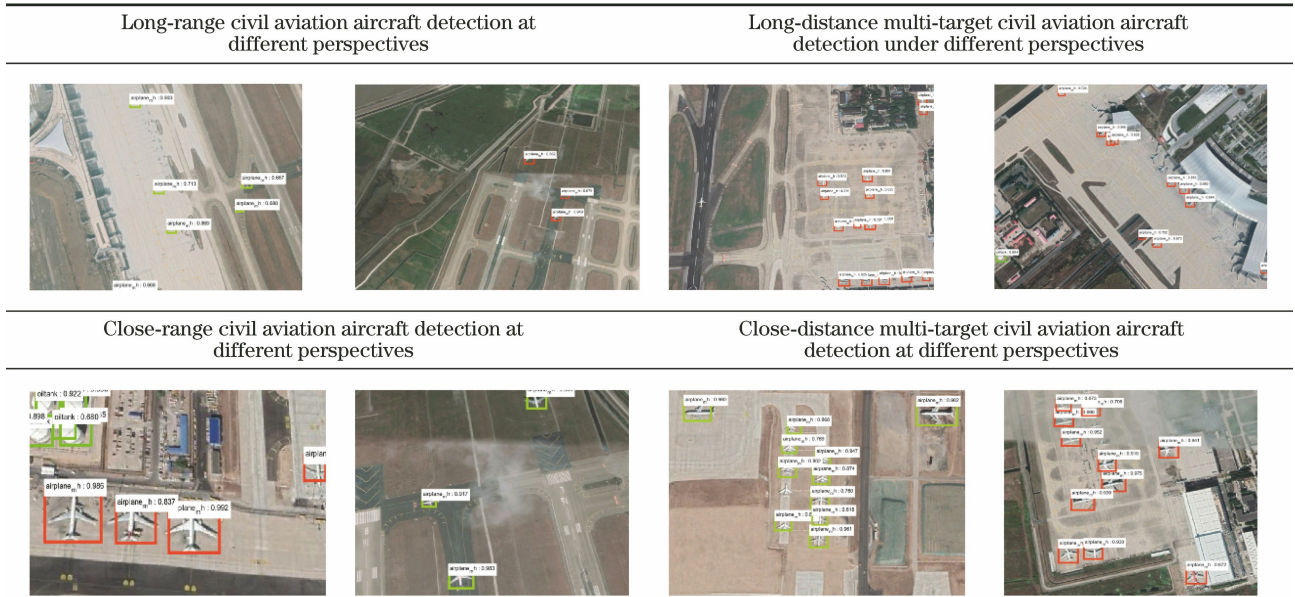


表 10 民航飞机的检测结果

Table 10 Civil aviation aircraft test results

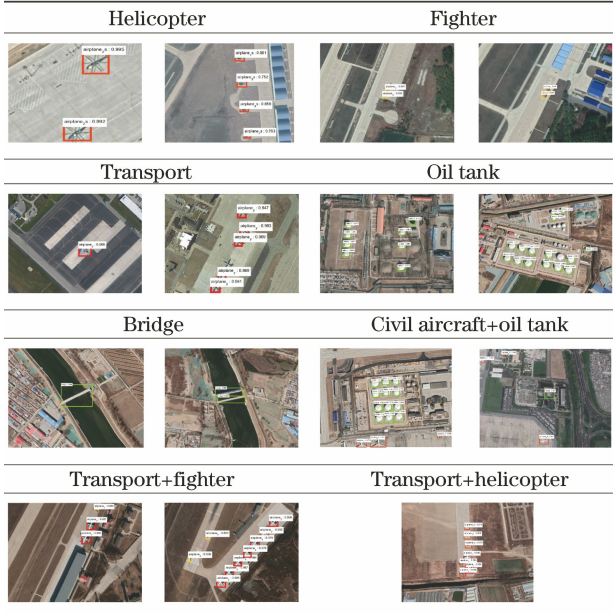


通过对以上实验结果分析,证明了改进的区域卷积神经网络同一模型对多类别、多尺度、多视角以及较为复杂背景下检测的可行性。针对大部分目标

的检测效果良好,拥有了较高的精确度,同时也暴露了复杂背景下小目标以及小样本目标的检测效果不强的问题,如图中存在超高空较为复杂背景下机场

表 11 多类别下的目标检测结果

Table 11 Target test results under multiple categories



的误检问题,不过较好的结果是通常误检测的目标置信度分数较低,再结合进一步的人工审阅,大多可以进行剔除。但是如表 11 所示,对于所占像素点很少的油罐目标,模型总是不能很好地将其检全,因此针对此方面的研究还需要进一步加强。此外还体现了数据库容量的重要性,构建更为广泛的数据库,将对模型的训练起到较大的改善作用。

3.4 对比仿真分析

使用高度一致的仿真环境以及网上公开的代码对原算法以及已有的飞机类和机场的检测方法进行对比实验。因为网上对其他目标的检测文献

较少且公开代码较少,在此未进行对比。检测结果如表 12、13 所示。

由表 12、13 可知,本文的改进算法除了相比原算法对复杂条件下机场、民航飞机、战斗机、运输机、直升机、油罐以及桥梁有更好的检测率外,与部分已有的民航飞机以及机场的检测方法相比,在平均精度方面均有一定提升,在检测时间方面也较为突出。文献[19]算法和文献[7]算法是人工设计特征结合卷积神经网络进行特征提取的经典的检测方法,文献[19]利用 BING 算法选取建议区域并利用 CNN 强大的提取特征能力,在一定程度上减少了区域建议的时间但是依旧长于区域建议网络,同时算法的检测精确率也低于本文算法。文献[7]算法依然存在复杂条件下人工设计特征的不足以及传统区域建议选取方法的局限性,因而平均检测率和检测效率均低于本文模型。Faster R-CNN 采用 CNN+RPN 结合的检测网络在很大程度上提高了检测效率,相比较而言,本文算法在进行复杂样本的筛选以及先验判决时增加了一些检测时间成本,但是从结果看,增量并不多且依旧有很高的效率,此外检测率提高了近 4%。综上所述,文献[19]与文献[7]算法都可以归类于人工设计特征与卷积神经网络的结合,是传统算法中检测飞机以及机场的典型算法。通过与传统经典检测算法的对比,突出了基于深度学习的区域卷积神经网络的端对端检测的速度优势以及精确度优势,同时与 Faster R-CNN 的对比体现了本文算法改进的有效性。

表 12 改进方法与原算法的各类目标检测结果对比

Table 12 Comparison of various target detection results between improved method and original algorithm

Method	AP / %							Mean time / s
	Airport	Civil aircraft	Helicopter	Fighter	Transport	Oil tank	Bridge	
Faster R-CNN	76.66	80.56	66.82	58.62	67.56	69.02	64.85	0.142
Proposed	80.84	84.82	70.10	62.14	71.01	73.59	68.73	0.145

表 13 不同检测方法的结果对比

Table 13 Comparison of results of different detection methods

Object	Method	AP / %	Mean time / s
Civil aircraft	Ref. [19]	76.73	6.87
	Faster R-CNN	80.56	0.142
	Proposed	84.82	0.145
	Ref. [7]	72.78	20.86
Airport	Faster R-CNN	76.66	0.142
	Proposed	80.84	0.145

4 结 论

本文构建了完备性更好、特点更加鲜明的数据集并对 Faster R-CNN 算法的局限性方面做出了相应改进优化,结果表明,优化算法在仅增加少许检测时间的基础上较大地提高了检测精度,在一定程度上改善了传统遥感图像检测方法对机场区域多种状态下多类目标检测鲁棒性不强的问题。此外,通过仿真得到的区域卷积神经网络目前对数据集的依赖性较强,对数据集中相应图像较少的目标进行尽可

能的扩充,以及对模型进行更深层次的改进,以进一步提高检测效果将是下一步研究的主要工作。

## 参 考 文 献

- [1] Schott J R. Remote sensing [M]. Oxford: Oxford University Press, 2007: 13-24.
- [2] Richards J A. Analysis of remotely sensed data: the formative decades and the future [J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3): 422-432.
- [3] Yao J, Zhang Z F. Semi-supervised learning based object detection in aerial imagery [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05), San Diego, CA, USA. IEEE, 2005: 1011-1016.
- [4] Zhao D P, Ma Y Y, Jiang Z G, et al. Multiresolution airport detection via hierarchical reinforcement learning saliency model[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(6): 2855-2866.
- [5] Tao C, Tan Y H, Cai H J, et al. Airport detection from large IKONOS images using clustered SIFT keypoints and region information[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(1): 128-132.
- [6] Zhu T H, Li Y H, Ye Q K, et al. Integrating saliency and ResNet for airport detection in large-size remote sensing images[C] // 2017 2nd International Conference on Image, Vision and Computing (ICIVC), June 2-4, 2017. Chengdu. IEEE, 2017: 20-25.
- [7] Zhang P, Niu X, Dou Y, et al. Airport detection on optical satellite images using deep convolutional neural networks [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1183-1187.
- [8] Chen X Y, Xiang S M, Liu C L, et al. Aircraft detection by deep belief nets [C] // 2013 2nd IAPR Asian Conference on Pattern Recognition, November 5-8, 2013. Naha, Japan. IEEE, 2013: 54-58.
- [9] Dai C K, Li Y. Aeroplane detection in static aerodrome based on Faster RCNN and multi-part model[J]. Journal of Computer Applications, 2017, 37(z2): 85-88.  
戴陈卡, 李毅. 基于 Faster RCNN 以及多部件结合的机场场面静态飞机检测 [J]. 计算机应用, 2017, 37(z2): 85-88.
- [10] Zhu M M, Xu Y L, Ma S P, et al. Airplane detection based on feature fusion and soft decision in remote sensing images[J]. Acta Optica Sinica, 2019, 39(2):0210001.
- [11] Shan Q W, Zheng X B, He X H, et al. Fast object detection and recognition algorithm based on improved multi-scale feature maps [J]. Laser & Optoelectronics Progress, 2019, 56(2):021002.  
单倩文, 郑新波, 何小海, 等. 基于改进多尺度特征图的目标快速检测与识别算法 [J]. 激光与光电子学进展, 2019, 56(2): 021002.
- [12] Zhang C, Chen Y. Object detection based on hard examples mining using residual network[J]. Laser & Optoelectronics Progress, 2018, 55(10): 101003.  
张超, 陈莹. 残差网络下基于困难样本挖掘的目标检测 [J]. 激光与光电子学进展, 2018, 55(10): 101003.
- [13] Sun X, Fu K, Wang H Q. High resolution remote sensing image understanding [M]. Beijing: Science Press, 2011: 13-25.  
孙显, 付琨, 王宏琦. 高分辨率遥感图像理解 [M]. 北京: 科学出版社, 2011: 13-25.
- [14] Richards J A, Jia X P. Remote sensing digital image analysis [M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999: 15-27.
- [15] 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.
- [16] Vicente S, Carreira J, Agapito L, et al. Reconstructing PASCAL VOC [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014. Columbus, OH, USA. IEEE, 2014: 320-327.
- [17] Chen F, Ren R L, van de Voorde T, et al. Fast automatic airport detection in remote sensing images using convolutional neural networks [J]. Remote Sensing, 2018, 10(3): 443.
- [18] Qassim H, Verma A, Feinzimer D. Compressed residual-VGG16 CNN model for big data places image recognition[C] // 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), January 8-10, 2018. Las Vegas, NV. IEEE, 2018: 440-446.
- [19] Wu H, Zhang H, Zhang J F, et al. Fast aircraft detection in satellite images based on convolutional neural networks [C] // 2015 IEEE International Conference on Image Processing (ICIP), September 27-30, 2015. Quebec City, QC, Canada. IEEE, 2015: 4210-4214.