第 40 卷 第 16 期
2020 年 8 月

光 学 学 报
Acta Optica Sinica

Vol. 40, No. 16
August, 2020

联合连续学习的残差网络遥感影像机场目标检测方法

李竺强[1*], 朱瑞飞[1,2], 马经宇[1], 孟祥玉[3], 王栋[1,2], 刘思言[1]

[1]长光卫星技术有限公司,吉林省卫星遥感应用技术重点实验室, 吉林 长春 130000;

[2]中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033;

[3]吉林省国土资源调查规划研究院, 吉林 长春 130061

摘要 在现有的高分辨率、大尺度目标遥感图像的检测中,传统方法由于提取特征手段单一、速度慢而无法快速并准确地从光学遥感影像中实现机场目标的识别。受人类视觉系统层次认知的启发,提出了一种适用于中高分辨率光学遥感影像的机场目标检测网络(CLRNet)。首先构建深度残差块,并将其作为特征提取网络;然后基于生成的样本核心集,采用连续学习方式从海量遥感数据中逐次迭代,精调机场检测模型;经过连续学习得到了鲁棒性强、遗忘度低的检测模型,该模型可以准确快速地从海量复杂背景下的光学遥感影像中识别出机场目标,而且对薄云遮挡以及卫星拍摄不全的机场有较好的识别效果。选取国产吉林一号卫星影像数据集进行测试,结果表明:所提方法的检测精度 mAP(IoU 不小于 0.5)可达 0.9613,每景的检测时间为 0.23 s。

关键词 遥感; 连续学习; 核心集; 机场检测; 残差卷积神经网络

中图分类号 TP751.1 **文献标志码** A **doi**: 10.3788/AOS202040.1628005

Airport Detection Method Combined with Continuous Learning of Residual-Based Network on Remote Sensing Image

Li Zhuqiang[1*], Zhu Ruifei[1,2], Ma Jingyu[1], Meng Xiangyu[3], Wang Dong[1,2], Liu Siyan[1]

[1] *Key Laboratory of Satellite Remote Sensing Application Technology of Jilin Province, Chang Guang Satellite Technology Co., Ltd., Changchun, Jilin* 130000, *China*;

[2] *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin* 130033, *China*;

[3] *Jilin Province Land Survey & Planning Institute, Changchun, Jilin* 130061, *China*

Abstract In the existing high-resolution and large-scale target remote sensing image object detection, the traditional method cannot achieve airport target recognition from optical remote sensing images quickly and accurately due to the single feature extraction and slow speed. Inspired by the hierarchical cognition of the human visual system, the continuous learning of residual-based convolution neural network (CLRNet) suitable for medium and high resolution optical remote sensing images is proposed. Firstly, the depth residual block is constructed as the feature extraction network. Secondly, the continuous learning method is used to fine tune the airport detection model from the massive remote sensing data. After continuous learning process, the model with strong robustness and low forgetting degree is obtained. The model can accurately and quickly identify airport from optical remote sensing images under massive and complex backgrounds. Our model has a better recognition effect for airports covered by thin clouds or incompletely captured by satellites. The domestic Jilin-1 satellite image dataset is selected for testing. Experiments show that the accuracy of the detection method mAP (IoU is not less than 0.5) can reach 0.9613, and the detection speed can reach 0.23 s per scene.

Key words remote sensing; continuous learning; core set; airport detection; residual convolution neural network

OCIS codes 280.4991; 100.2000; 100.3008

收稿日期: 2020-03-28; **修回日期**: 2020-05-12; **录用日期**: 2020-05-18

基金项目: 国家重点研发计划重点专项(2018YFB1004605);吉林省重点科技研发项目(20180201109GX)

* **E-mail**: skybelongtous@foxmail.com

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1 引　言

随着高分专项的实施和商业化卫星遥感服务的兴起，高分辨率卫星遥感数据日益丰富，应用研究逐步深入[1]。遥感卫星图像系统和无人机对地观测技术的成熟与发展使得目前的高分数据图像数量不断增多，分辨率不断提升，数据的实时性越来越高，从而对目标的检测精度和效率提出了挑战。在现有的大数据量高分图像处理方法中，传统方法由于提取特征的手段单一、速度慢，无法在影像质量低时实现关键目标的检测与识别。如何从海量的非结构化数据中，挖掘到有价值的目标信息成为学术界与工业界面临的难题。对于各种类型的飞行器来说，机场是非常重要的承载设施，快速并准确地从光学遥感影像中识别机场在军事上具有重要意义。

目前，通过遥感影像检测机场从机理上来讲有三种方法：通过基于边缘直线的方法来确定机场跑道；基于机场纹理结构的显著性特征进行识别；基于深度学习的目标检测，即通过从海量样本中构建深度卷积特征进行识别。

基于边缘直线的方法主要是利用飞机跑道明显的直线特征来反映目标的性质。Wang 等[2]首先利用 Hough 变化[3]判断机场是否存在，然后采用改进的基于图的视觉显著性注意力模型提取机场的候选区域。然而由于机场形状、周围背景环境以及薄云等的影响，检测直线特征的模式具有一定的局限性，会将形状与机场跑道类似的高速公路误检为机场。基于目标的显著性结构进行识别的方法主要通过视觉显著性机制实现目标的识别[4]。视觉感知层会关注结构特征明显或连通域较大的区域，并会刺激脑神经形成对目标的固有印象，这一过程是一种知识驱动型显著性机制[5]。此类方法一般能够很好地估计目标的显著性，但往往容易突出边缘特征，而场景中显著性最高的区域有可能不是机场目标，因此很难在复杂的中高分辨率遥感影像场景中取得良好的识别效果。

随着深度学习浪潮的到来，以深度卷积运算为内核、基于图形处理器（GPU）的并行技术在图像识别领域被广泛应用。余东行等[6]和刘媛等[7]提出了一种显著性特征与卷积神经网络相融合的遥感影像机场检测算法。Chen 等[8]利用 Faster R-CNN (region-based convolutional network) 方式实现了机场目标的识别。为了提升检测效率，Zhang 等[9]在原有的卷积神经网络中，用一种线段分割的弱语义规则来替代原有的区域生成网络（RPN）；虽然效率得到了提升，但误检目标也增多，整体检测精度为 88.8%。Cai 等[10]和朱明明等[11]基于改进区域建议网络和损失函数的方法，实现了复杂背景下的机场检测，但由于样本量不足，出现了检测框不完整或一个机场多个包围框的情况。张作省等[12]将直线检测与深度卷积神经网络影像分类相结合，先利用 Canny 算子提取灰度图的边缘信息，通过判断平行线是否存来实现初步筛选，然后再次利用深度卷积神经网络进行判别，以得到相应区域的识别概率值。这种方法虽然可以提高识别效率，但得到的定位准确率较低，为 87.5%。Kirkpatrick 等[13]提出了弹性权重巩固（EWC）梯度优化方法，该方法不仅可使神经网络模型具备多次任务连续学习的能力，还可使模型在学习增量训练数据的同时不遗忘之前模型的信息[14]。

针对复杂环境下遥感影像中机场目标检测精度低的问题，本文提出了一种基于联合连续学习的残差块网络模型 CLRNet。为解决少量样本训练梯度爆炸的问题，本文采用一种基于尺度变换的增量学习进行数据扩充。同时，对增量后的样本构建核心集[15]，使训练后的模型具有低遗忘度，并利用连续学习模式对残差块网络进行权重更新。为提升模型在复杂环境下对机场的识别效果，在特征提取阶段采用残差块卷积神经网络 ResNet 101 进行特征图提取。此外，针对机场特点合理改进 RPN 锚点（anchor）缩放尺度的层次性，以减少训练计算量。为了验证检测方法有效性，选取国产卫星吉林一号数据集，对目前流行的检测算法进行纵向对比。实验结果表明，本文提出的检测网络模型表现出了更加优越的识别性能。

2 实验方法

图 1 为本文提出的联合连续学习的残差网络机场目标检测技术路线图。流程如下所示：

1）基于尺度变换的预处理方法以适宜分辨率的影像作为网络输入。为提升后续模型的泛化能力和鲁棒性，采用对样本进行旋转、平移、色彩空间变换、对比度变换以及添加高斯噪声等的方法扩增数据量。同时，针对持续性数据下目标实时更新慢的问题，构建训练样本核心集。

2）自然图像中的目标与遥感影像中的目标在基本视觉模块（边、角特征）中有一定的相似性，通过迁移学习的方法来优化深度卷积网络。利用残差网

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

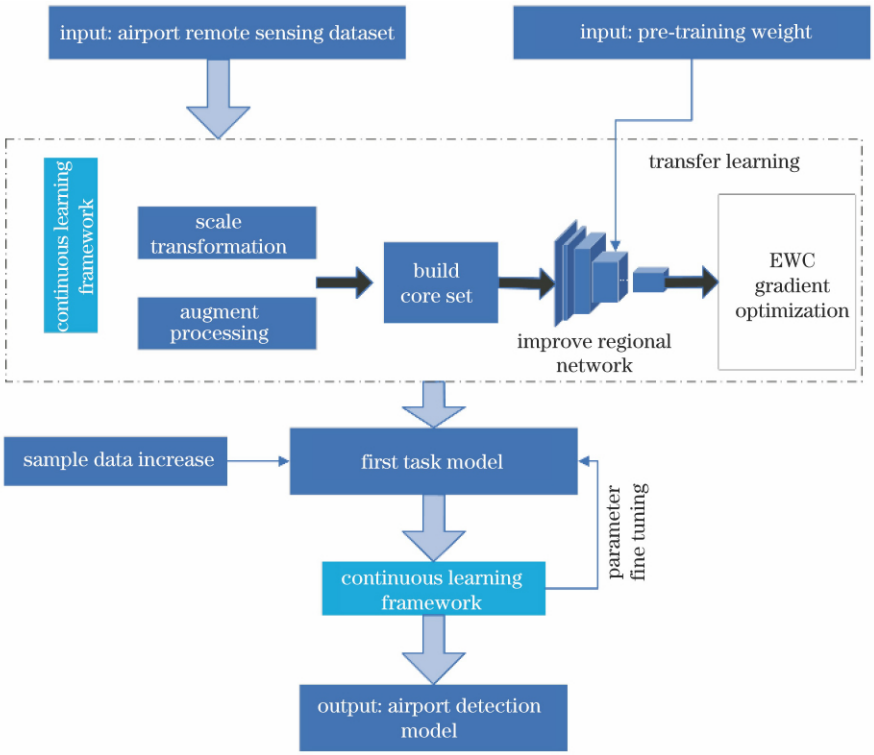

图 1　联合连续学习的残差网络机场检测流程图

Fig. 1　Flow chart of airport detection method combined with continuous learning of residual-based network

络在 ImageNet 数据集中的权重参数对遥感目标识别骨干特征网络进行预训练，初始化区域生成网络每层的参数，保留卷积层的权重。

3）在训练样本核心集上利用步骤 2）中迁移学习得到的初始化卷积层权重进行首次模型训练，得到当前最佳的卷积神经目标检测模型 CLRNet 1st stage，该模型可在参数空间达到局部最优。网络训练优化函数使用 EWC 梯度优化函数，根据网络参数的重要程度对其施加不同程度的约束，最终得到的参数配置会同时在首次模型 CLRNet 1st stage 和最终模型 CLRNet 2nd stage 中表现出良好的性能。

4）当影像数据更新后，只需要在小规模相关数据集中采用连续学习网络对首次模型 CLRNet 1st stage 进行微调训练，就可以得到最终的目标识别模型 CLRNet 2nd stage。在少量样本更新下提升训练速度，可以缓解海量遥感影像样本数据因持续更新而导致模型更新滞后的问题。

2.1　基于尺度变换的数据增量预处理

为更快速地从卫星遥感影像中得到目标位置，本文构建了“数据-尺度化-像素-目标”层次认知检测模型。将具有合适分辨率的影像作为网络输入，这样在训练过程中就不会因为影像输入尺寸过大而使得卷积层过多，从而导致模型参数难以收敛。

$$\max\{I_{\mathrm{w}}, I_{\mathrm{h}}\} \cdot r_1 \cdot \rho \approx N_{\mathrm{input}}, \tag{1}$$

$$r_{\mathrm{obj}} = \frac{l_{\mathrm{obj}}}{I_{\mathrm{obj}}}, \tag{2}$$

$$I_{\mathrm{obj}} \in [7, N_{\mathrm{input}}), \tag{3}$$

式中：I_{w} 与 I_{h} 为卫星推扫多光谱影像单景像元的行数和列数；r_1 为原始影像的分辨率；ρ 为降尺度系数；N_{input} 为最佳输入影像的大小；l_{obj} 为机场长度；I_{obj} 为尺度变化后机场目标的像元大小；r_{obj} 为尺度变化后机场目标的分辨率。由航空民航标准 MH 5001—2013《民用机场飞行区技术标准》可知，机场跑道长度一般大于 3 km，故可认为机场长度 $l_{\mathrm{obj}} \in [3\ \mathrm{km}, 10\ \mathrm{km}]$。由于检测网络框架卷积核大小的限制，目标区域的大小应该大于 I_{obj} 像元（7×7）。由（2）、（3）式可知 r_{obj} 与 ρ 的关系为 $r_{\mathrm{obj}} = \frac{1}{\rho}$，尺度变化后机场目标的分辨率 r_{obj} 为一个范围值，为 14～23 m/pixel。

对尺度变化后的样本数据进行增量处理，包括影像旋转、翻折、平移、色彩对比度变化以及随机添加高斯噪声，目的是提升后续网络模型的泛化能力和鲁棒性。

2.2　连续学习核心集构建

在大规模遥感数据的背景下，目标检测模型实

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

时更新的代价较大。一方面，影像数据持续增加造成重新训练时间代价上升；另一方面，样本数据量更新导致模型无法在给定的时间内学习到代表性特征，从而导致检测精度下降。针对持续性数据下目标实时更新慢的问题，本文选用增量后的样本数据构建训练核心集，并对样本进行非均匀采样来保证样本尺度特征的差异性大，同时减小简单机场目标样本在训练中的比例。

训练核心集是从整个增量后的影像数据集中非均匀采样形成的一个子集，用所选子集训练得到的网络模型与用整个数据集得到的模型的性能一致。平衡数据集中复杂样本与简单样本的比例，可以保证特征的多样性，同时可以减少训练时间和存储空间。

为使所选的样本影像具有代表性，本文采用了一种基于 k-Means 非均匀特征采样的方法。对于影像背景复杂的样本，该方法可使其被采样的概率增大，从而可以解决复杂样本与简单样本不平衡的问题。计算方法如下：基于冻结全连接层的 ResNet 预训练网络对增量后的样本数据进行特征提取，通过 k-Means 在特征空间进行 k 个中心聚类，本文选取的两个聚类中心分别为复杂背景样本 C_{complex} 与简单背景样本 C_{simple} 的特征中心。设增量后影像集合 R_{aug} 包含 n 景影像，选取 m 景影像作为该数据的训练核心集。集合 T_{train} 表示从集合 R_{aug} 中抽取到的影像组成的集合。计算步骤如下：

1）首先在两个特征中心 $\{C_{\text{complex}},\ C_{\text{simple}}\}$ 附近的样本中随机选取两个点 $q_{\text{s}}^{(1)}$、$q_{\text{c}}^{(1)}$ 作为特征中心，并将这两个点放入集合 T_{train} 中，将 $q_{\text{s}}^{(1)}$、$q_{\text{c}}^{(1)}$ 从 R_{aug} 中移除。

2）分别查找特征中心 $q_{\text{s}}^{(1)}$、$q_{\text{c}}^{(1)}$ 的 K 个邻近点，$K\geqslant 4$。

3）若点 $q^{(i)}$ 满足特征空间 $\frac{\mu}{2}\leqslant d^2(o,q^{(i)})\leqslant\mu$，则将点 $q^{(i)}$ 放入集合 T_{train} 中，并从 R_{aug} 中移除点 $q^{(i)}$。其中，o 为特征中心，μ 为平均特征距离，$q^{(i)}$ 为 2）中的邻近点。由于复杂样本特征的差异较大，故而对 C_{complex} 中心采样可以得到更多的样本点。

4）更新特征中心 o，特征逐渐变化的点（即满足条件的邻近点）被选为新中心点，用它代替初始 $q_{\text{s}}^{(1)}$、$q_{\text{c}}^{(1)}$ 再次进行 K 邻近选取新点，并判断该点是否满足 3）；

5）按步骤 4）进行迭代，直到从集合 R_{aug} 中采样到 m 景影像为止，最终得到由集合 T_{train} 构造的核心集，$T_{\text{train}}=\{q_{\text{s}}^{(1)},q_{\text{c}}^{(1)},\cdots,q_{\text{s}}^{(x)},\ q_{\text{c}}^{(y)}\mid x+y\leqslant m\}$。

遥感影像数据是持续拍摄积累的，通常不能一次性构建核心集。下面通过二叉树方法构造机场目标核心集。如图 2 所示，$d_1\sim d_{16}$ 为持续时间序列的数据集影像。先基于数据集 $d_1\sim d_4$，利用 k-Means 非均匀特征采样方法构造核心集 C_1；类似地，再基于数据集 $d_5\sim d_8$ 构造核心集 C_2；然后合并核心集 C_1 与 C_2 得到核心集 C_3。同时，C_3 亦为数据集 $d_1\sim d_8$ 的核心集。

采用图 2 所示的二叉树方式构造核心集 $C_1\sim C_7$，C_7 为最终数据集 $d_1\sim d_{16}$ 的核心集。基于二叉树方式构造核心集可以解决因前期数据不足而使训练停滞的问题；同时，数据汇聚后也便于模型的更迭。

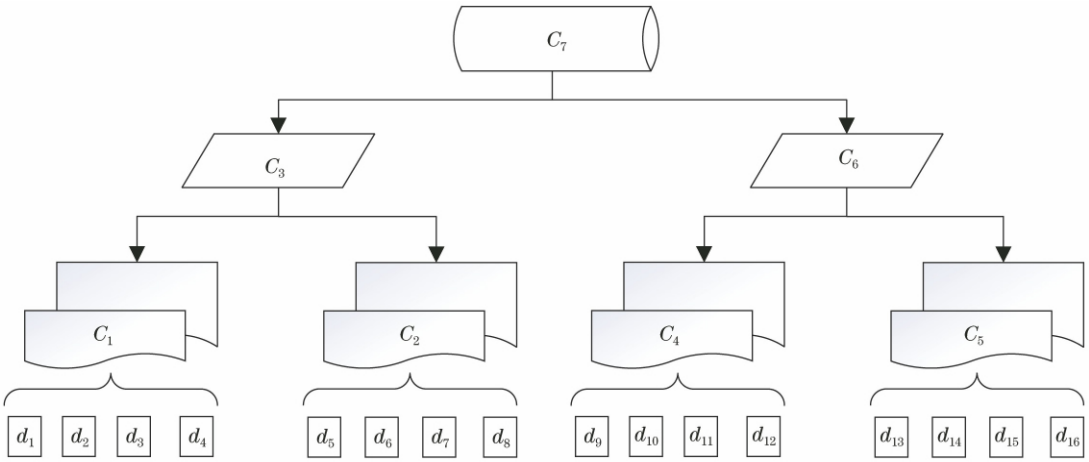

图 2　基于二叉树方法构造机场目标核心集

Fig. 2　Constructing airport object core set by binary tree method

2.3　网络结构

结合残差块网络的机场目标检测网络结构如图 3 所示。首先以深度残差块网络 ResNet101 作为骨干网络对图像进行特征提取，接下来用 RPN

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

在上一步输出的特征中预测疑似机场区域，最后联合检测网络的感兴趣区(ROI)池化与非极大值抑制(NMS)对机场目标进行分类和边界位置回归。

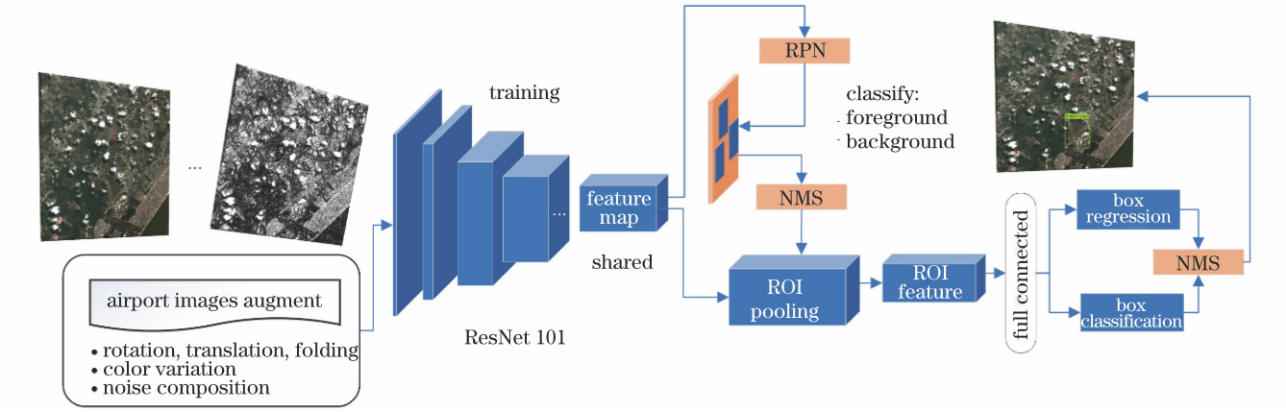

图 3　结合残差块网络的机场目标检测网络结构

Fig. 3　Airport object detection network structure combined with residual block network

区域生成网络是一个联合残差块的全卷积神经网络，其输入为残差块特征提取网络得到的特征图，输出为一系列的矩形目标候选区。后续检测网络与区域生成网络共享卷积特征图，可以大幅减小计算量，并使建议框的初次识别准确率更高。

改进的 RPN 如图 4 所示，RPN 在 ResNet 101 残差块网络输出特征图上进行滑窗移动卷积。滑窗卷积核为 3×3 大小的视窗，每个视窗可以同时预测 n 个目标候选区，这 n 个候选区都与这个视窗存在关联。每个视窗点(anchor)都有着对应的尺度和比例。卷积特征图中的每一个点都是一个 anchor 中心，对于 14×14 大小的卷积特征图，存在 196 个锚点。每个视窗被映射为一个 256 维的特征向量，并通过反向传播网络回归目标的分类与边界位置。目标分类回归网络输出的每个锚点前景与背景的概率，边界回归网络输出其平面方向的偏移值。为了加速预测过程以及应对遥感图像中的多尺度机场目标，本文考虑了两种尺寸(128×128 与 512×512)和三种长宽比(1∶1、1∶2、2∶1)，以减少建议框的数量。

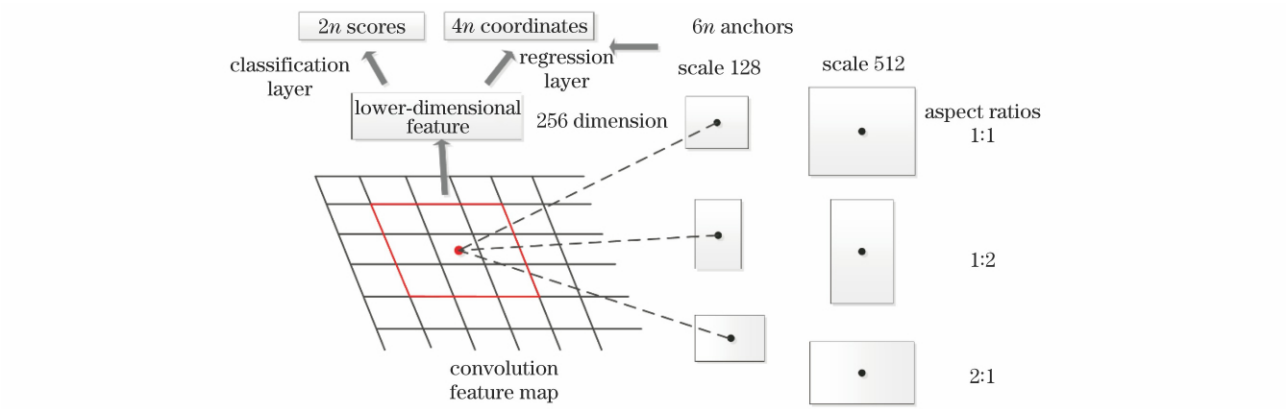

图 4　针对机场目标改进 RPN 的 anchor 缩放尺度

Fig. 4　Improve the scale of anchor in RPN for airport object

区域生成网络的训练过程是端到端的，使用的优化方法是反向传播和 EWC 梯度优化函数。损失函数是分类误差和回归误差的联合损失，即

$$L(\{p_i\}+\{t_i\})=\frac{1}{N_{\mathrm{cls}}}\cdot\sum_i L_{\mathrm{cls}}(p_i,p_i^*)+\frac{\lambda}{N_{\mathrm{reg}}}\cdot\sum_i p_i^*\cdot L_{\mathrm{reg}}(t_i,t_i^*),\tag{4}$$

式中：i 表示第 i 个锚点；p_i 表示第 i 个锚点中目标的预测概率；p_i^* 表示第 i 个锚点为真值的标记，若锚点的真实标签为正样本，则 $p_i^*=1$，反之为 0；t_i 表示第 i 个锚点映射候选边框的位置；t_i^* 表示第 i 个锚点真实目标边框的位置；λ 表示边框定位损失权重；N_{cls} 为训练时批量样本的数量；L_{cls} 为分类损失函数；N_{reg} 为批量样本锚点的数量；$L(\cdot)$ 为分类误差和回归误差的联合损失；$L_{\mathrm{reg}}(\cdot)$ 为边框回归损

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

失。在训练时，正样本为与任意边框、真实边框的交并比(IoU)大于 0.65 的候选框，负样本为与所有真实边界的 IoU 小于 0.35 的候选框。

区域生成网络会生成目标建议框，如图 4 所示。采用非极大值抑制滤掉重叠率 IoU 大于 0.65 的框后，连接目标检测网络，利用感兴趣区域池化层从完整的特征图上直接得到各建议框对应的特征矢量，经过 3×3 的卷积层与全连接层后输出维度为[$B\times128$，2048]的特征向量(B 表示 batch size)，该特征向量参与目标类别和边界框的预测。

分类层输出的是每个感兴趣区域在 $k_{\text{class}}+1$ 个类别上的概率分布 $p=(p_0,\cdots,p_{k_{\text{class}}})$。边界回归层输出的是边界框回归的参数，$t^{k_{\text{class}}}=(t_x^{k_{\text{class}}},t_y^{k_{\text{class}}},t_w^{k_{\text{class}}},t_h^{k_{\text{class}}})$，其中 k_{class} 表示目标类别(本文取 $k_{\text{class}}=2$，一类是机场目标，另一类为纹理结构相似的负样本)。这两个输出层通过联合损失函数进行训练，该函数的表达式为

$$L(p,u,t^u,v)=L_{\text{cls}}(p,u)+\lambda[u\geqslant1]\cdot L_{\text{loc}}(t^u,g),\tag{5}$$

式中：$L_{\text{cls}}(p,u)=-\ln(p_u)$ 是真实类别 u 的对数损失；$L_{\text{loc}}(\cdot)$ 为目标边框损失；t^u 为类别 u 的预测边框；g 为类别 u 的真实边框。设类别 u 的真实边框 $g=(g_x,g_y,g_w,g_h)$，类别 u 的预测边框 $t^u=(t_x^u,t_y^u,t_w^u,t_h^u)$，其中各参数的计算公式为

$$t_x=\frac{B_x-Y_x}{Y_w},\tag{6}$$

$$t_y=\frac{B_y-Y_y}{Y_h},\tag{7}$$

$$t_w=\ln\left(\frac{B_w}{Y_w}\right),\tag{8}$$

$$t_h=\ln\left(\frac{B_h}{Y_h}\right),\tag{9}$$

式中：B_x 和 B_y 表示真实目标的中心坐标；B_w、B_h 分别表示边框的宽和高；Y_x 和 Y_y 表示候选区域预测的中心坐标；Y_w、Y_h 分别表示候选区域的宽和高。定义边框回归层损失函数表达式为

$$L_{\text{loc}}(t^u,g)=\sum_{i\in\{x,y,w,h\}}S_{L_i}(t_i^u-g_i),\tag{10}$$

其中，

$$S_{L_i}(x)=\begin{cases}|x|-0.5, & |x|\geqslant1\\ 0.5x^2, & \text{otherwise}\end{cases}。\tag{11}$$

图 5 为采用连续学习训练好的检测网络对目标影像进行特征图提取过程的可视化描述。图 5(a)为一幅包含机场目标的真彩色影像；图 5(b)为采用残差网络中的 64 个卷积核进行卷积得到的特征图；图 5(c)为采用残差网络中的 1024 个卷积核得到特征图；图 5(d)为最终分类回归与边框回归得到的特征热力图，可以看出检测网络对线性跑道特征及其附属机场绿地有较高的敏感性。

2.4　训练策略

共享卷积层的权重用 ResNet 101 预训练网络进行初始化，其余层的权重用均值为 0、标准差为 0.01的高斯分布函数进行随机初始化。网络梯度优化函数为 EWC，其基本学习率为0.007，动量为0.9，权重衰减系数为 0.5，分别在第 100、第 200 个 epoch 处衰减。模型保护权重 η 为 1.5，Fisher 信息矩阵采样数量为 2048。为了实现区域生成网络和目标检测网络共享卷积层，采用交替优化策略训练整个网络，具体步骤如下：

1) 利用预训练模型 ResNet 101 初始化骨干特征提取网络，并独立训练一个区域生成网络。

2) 进行批处理训练，每一批均是从 B 张图片中获取锚点框(B 为 2)，且要求正样本和负样本的数量相同，$B\times128$ 个正、负样本。

3) 将步骤 2)中区域生成网络生成的候选框进行非极大值抑制过滤后的目标框作为输入，对目标检测网络进行初始化，区域生成网络与目标检测网络的参数完全不共享。

4) 使用步骤 3)中的目标检测网络初始化一个新的区域生成网络，将区域生成网络、目标检测共享卷积层的 learning rate 设为 0，仅更新区域生成网络回归层；重新训练后，两个网络公共卷积层的参数完全共享。

5) 固定共享卷积层，加入目标检测感兴趣池化与全连接层，形成统一的网络并继续训练，完成目标检测网络类别与边界回归层的精调；最终得到首次机场检测网络模型 CLRNet 1st stage。

6) 当更少的数据更新时，只需要利用小规模相关数据集对模型 CLRNet 1st stage 进行连续学习网络微调训练，就可以得到最终的目标识别模型 CLRNet 2nd stage。在少量样本下，训练速度更新得非常快，这有助于缓解海量遥感影像样本数据持续更新而模型更新滞后的问题。此框架网络支持新任务优化训练以及其他典型目标的检测。

如图 6 所示，在训练时设置 B 为 2，训练 350 个 epoch。监控训练过程中验证集的评价指标如下：召回率(r)、精度(p)、F_1 指数、预测框中心点(x,y)的损失值、预测框尺寸(w,h)的损失值、预测目标类别

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

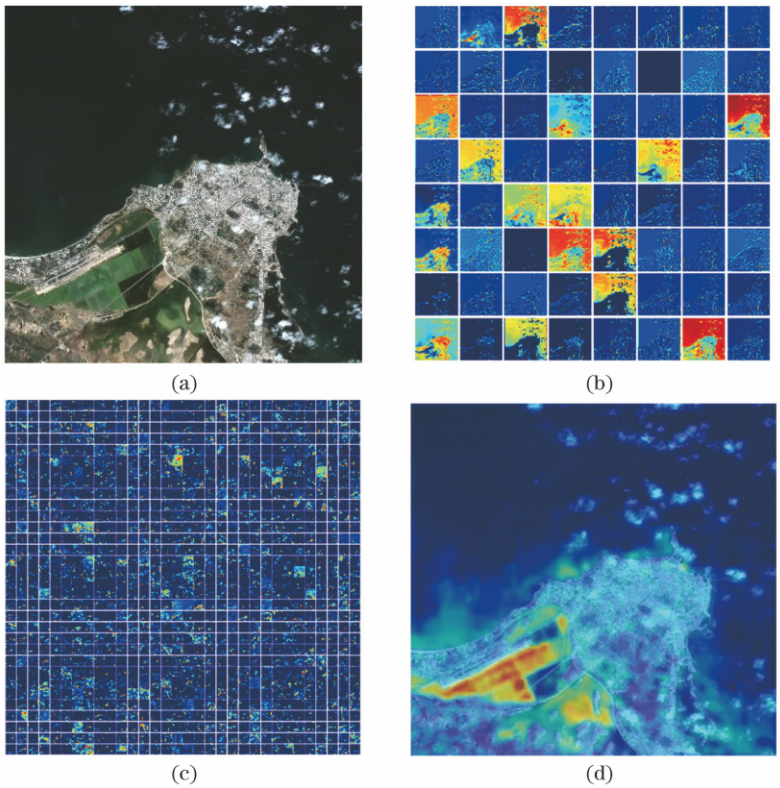

图 5　残差网络对影像中关注目标的特征图可视化。(a) 原始真彩色影像；(b) 64 个卷积核得到的特征图；(c) 1024 个卷积核得到的特征图；(d) 分类回归与边框回归得到的特征热力图

Fig. 5　Visualization of the feature map of the target in the image by the residual network. (a) Original true color image; (b) feature map from 64 convolution kernels; (c) feature map from 1024 convolution kernels; (d) feature heat map obtained from classification regression and border regression

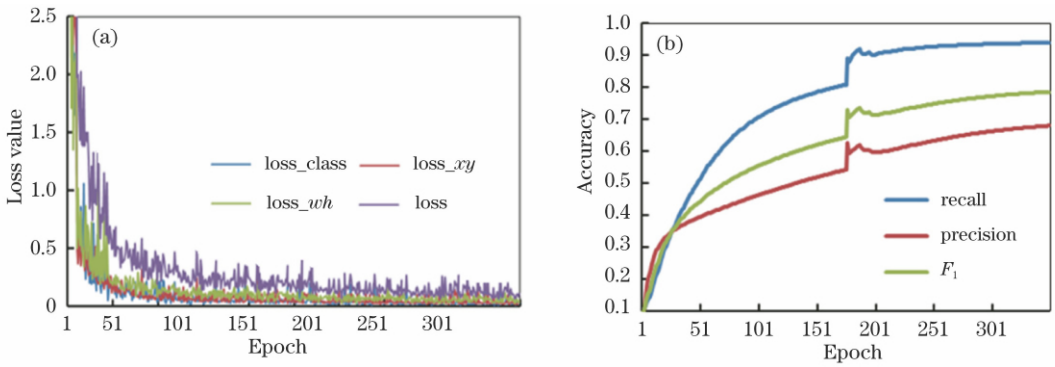

图 6　验证集损失函数的下降曲线与精度测试。(a)损失值下降曲线；(b)精度测试曲线

Fig. 6　Verification set loss function decline curve and accuracy test. (a) Loss value decline curve; (b) accuracy test curve

class 的损失值以及总体损失值。图 6(a)是验证集各部分损失随着训练过程的变化曲线，图 6(b)是验证集的召回率、精度和 F_1 指数随着训练过程的变化曲线。可以看出，验证集在第 250 个 epoch 训练后测试精度曲线处于平稳，损失值处于局部波动状态，此时基本可以使模型参数达到全局最优。

3　实验与结果分析

3.1　遥感数据集

数据集增量后共有 3068 幅遥感图像，将其中的 2147 幅作为训练图像，460 幅作为验证图像，461 幅作为测试图像。训练、验证与测试数据的比例大致为 70%∶15%∶15%。数据来源主要为吉林一号卫星星座的多光谱影像，主要使用 Jilin-101A、Jilin-

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

1SP 以及 Jilin-GF03 卫星的影像，空间分辨率为 2.88～4.24 m。详细的数据参数介绍如表 1 所示。训练与测试环境为 Ubuntu-16.04 系统，3.50 GHz Intel(R)Core(TM)i7-5930K CPU，内存为 64 GB，使用 GTX Titan X GPU@12 GB 进行加速运算。

表 1　吉林一号卫星星座的机场目标遥感数据集

Table 1　Airport object remote sensing dataset of Jilin-1

Year	Original panchromatic resolution /m	Original multispectral resolution /m	Network input resolution /m	Detected number /scene	Total images /scene
2016	0.72	2.88	14.4	2387	63645
2017	0.72	2.88	14.4	1206	47863
2018	0.72—0.92	2.88—3.68	14.4—18.4	4401	172311
2019	0.72—1.06	2.88—4.24	14.4—22.0	4200	191311

3.2　网络结构与参数设置

整个网络如表 2 所示，在残差网络特征提取阶段使用连续卷积核尺寸为 3×3 和 1×1 的卷积层与一些 shotcut 连接。网络相比 ResNet 50 层次更加丰富，相比 DenseNet 与 ResNet-152 具有更高的提取效率。在表 1 中，K 为普通卷积核尺寸，K_{rb1} 与 K_{rb2} 为 Residual block 中的卷积核尺寸，s 为卷积步长，p 为卷积操作前的影像填充大小。

在 RPN 候选框提取阶段，进行如下设置：anchor scales 取值为 9 和 36；anchor ratios 取值为 0.5、1、2；设置三种长宽比（即 1∶1、1∶2和 2∶1）；最大候选框个数为 350；nms 阈值为 0.65；边框定位损失权重（localization loss weight）为 1.5；正负样本损失权重（objectness loss weight）为 1.0。

在检测网络回归阶段，进行如下设置：正负样本比例（stage balance fraction）为 1∶2；为了在面对未知数据时也能保持训练时的正确率，加入 L2 正则化，初始权值设置为 0.0001；边框定位损失权重（localization loss weight）为 2.0；类别样本损失权重（classification loss weight）为 1.0。

表 2　结合残差块网络的机场目标检测网络结构表

Table 2　Airport object detection network structure combined with residual block network

Module	Kernels′ number	Convolution parameter	Input	Output
Conv 1	64	K:7×7, s:2, p:3	896×896×3	448×448×64
MaxPooling 1	64	K:3×3, s:2, p:1	448×448×64	224×224×64
Conv 2	64	K:1×1, s:1	224×224×64	224×224×64
Conv 3	192	K:3×3, s:2, p:1	224×224×64	112×112×192
MaxPooling 2	192	K:3×3, s:2, p:1	112×112×192	56×56×192
Conv 4	192	K:3×3, s:1, p:1	56×56×192	56×56×192
MaxPooling 3	192	K:3×3, s:2, p:1	56×56×192	28×28×192
Residual block 1	256	K_{rb1}:3×3, K_{rb2}:1×1	28×28×192	28×28×256
Residual block 2	320	K_{rb1}:3×3, K_{rb2}:1×1	28×28×256	28×28×320
Residual block 3(×3)	576	K_{rb1}:3×3, K_{rb2}:1×1	28×28×320	14×14×512
RPN Conv	256	K:3×3, s:1, p:1	14×14×512	14×14×256
ROI Pooling	-	-	14×14×256	4096×1
Fully connected layer	-	-	4096×1	2048×1

3.3　测试结果与分析

3.3.1　可视化定性分析

下面基于吉林一号卫星星座机场数据测试集可视化展示 CLRNet 网络对不同类型、背景环境下机场目标的检测进行分析。

图 7 所示为基于 CLRNet 1st stage 网络对不同背景环境下（丘陵、荒漠、海岛、港口）遥感影像中机场目标的测试结果。可以看出，检测边框可以很好地包围机场边界。在初始任务 1st stage 中，由于首次任务网络对多样化背景很难学习到最优参数，故而造成了背景纹理与机场相似的目标的误检，如图 8 所示。

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

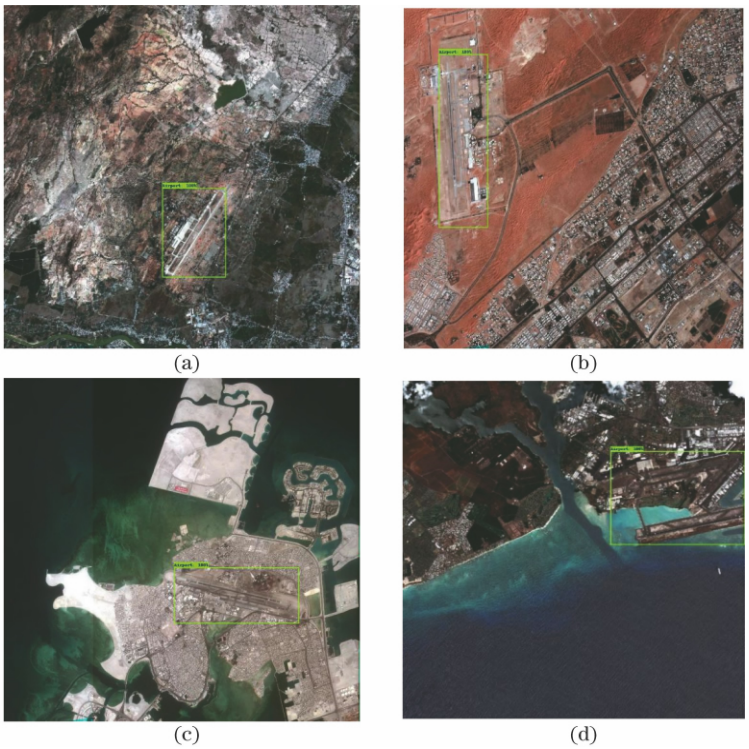

图 7　不同背景环境下遥感影像中机场目标的识别结果。(a) **丘陵区域环境下的测试结果**;(b) **荒漠环境下的测试结果**;(c) **海岛环境下的测试结果**;(d) **港口环境下的测试结果**

Fig. 7　Airport object detection results in remote sensing images under different background environments. (a) Test result in hilly area environment; (b) test result in desert environment; (c) test result in an island environment; (d) test result in port environment

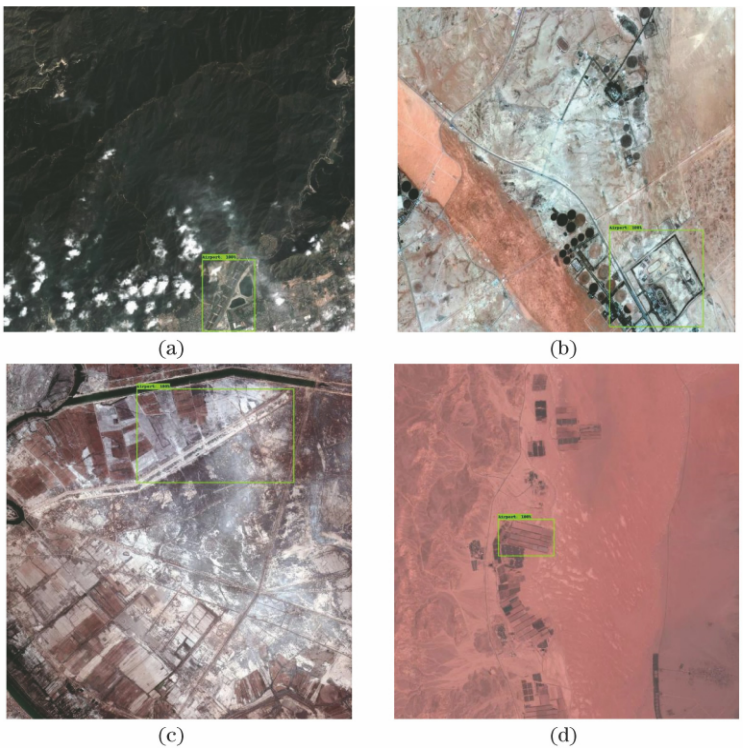

图 8　背景纹理与机场目标相似造成的误分情况。(a) **桥类设施**;(b) **工业园区**;(c) **公路**;(d)**结构化实验田**

Fig. 8　Misclassification caused by background texture similar to airport objects. (a) Bridge facilities; (b) industrial park; (c) highway; (d) structured experimental field

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

图 8 展示了机场目标的误检情况，其中的一个误检原因是桥类设施等目标与真实目标纹理结构比较相似，导致区别性特征表达不明显。图 8(c)中的一段公路与机场跑道比较相似，造成分类器识别目标产生混淆，将公路误检测成机场。另一个误检原因是目标与背景比较相似，目标的边、角特征使得其显著性增强。如图 8(d)所示，在一片荒漠中，将结构化的实验田错检成机场，造成虚警率上升。

如图 9 所示，通过连续学习 CLRNet 2nd stage 后，目标识别在干扰环境条件下表现优异。图 9(a)为海岸线附近一个机场的识别结果，由于云雾影响，整幅影像的厚云量约为 15%，薄云量约为 40%左右，机场西南角被遮挡。由于拍摄轨道问题，机场恰好被“切分”，如图 9(b)所示。图 9(c)为在同幅影像中存在面积差异较大的两个机场情况下目标的识别情况。图 9(d)是在比一般民用机场更小的小型机场干扰下目标的识别结果。

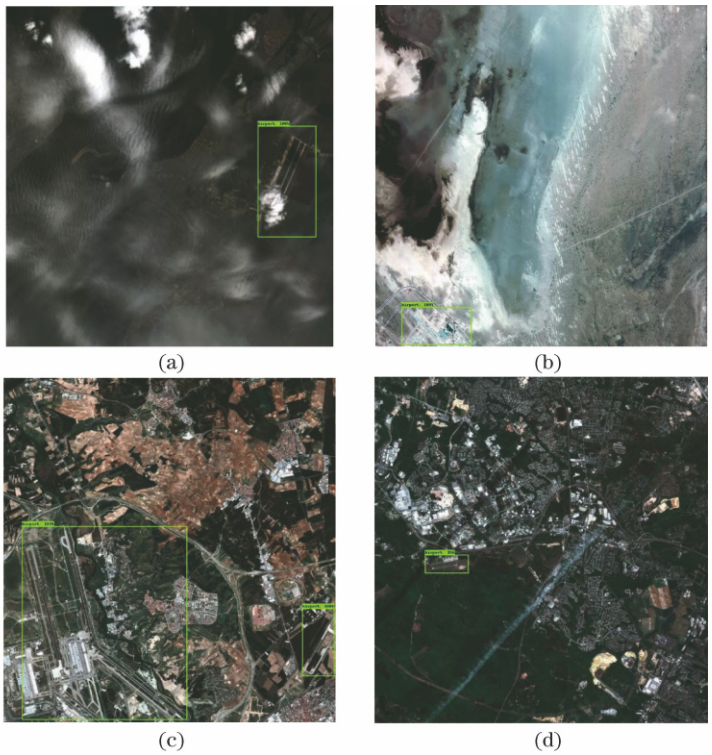

图 9　干扰条件下机场目标的识别结果。(a)云雾干扰；(b)推扫残缺；(c)同幅影像下目标尺度差异较大；(d)小型机场

Fig. 9　Airport object detection results under interference conditions. (a) Cloud interference; (b) sweeping incomplete; (c) large difference in object scale in the same image; (d) small airport

图 10 展示了连续学习模式下机场识别结果的对比，图 10(a)、(c)为 CLRNet 1st stage 的识别结果。从图 10(a)中可以看出，在首次模型学习中，将疑似机场的工业厂房误检出来。经过连续学习 CLRNet 2nd stage 后，该部分被检测为负样本。由图 10(c)可以看出，在积雪环境覆盖下，前景因与背景目标相似度低而被误检测。同样，经过连续学习 CLRNet 2nd stage 后，此类问题得以解决。

3.3.2　定量评价

采用精度 p、召回率 r、mAP^{50} 和 IoU 来评价检测准确率，在预测得到的目标与真实框的 IoU 大于 0.5 并且置信度大于 0.8 的基础上进行定量评价。本文使用的 IoU 阈值为 0.5，置信度阈值为 0.8。p 的计算公式为

$$p=\frac{T_{\mathrm{P}}}{T_{\mathrm{P}}+F_{\mathrm{P}}}=\frac{T_{\mathrm{P}}}{N},\tag{12}$$

式中：T_{P} 表示真正例；F_{P} 表示假正例；N 表示检测出来的目标数目。精度反映检测结果中检测正确的比例。召回率的计算公式为

$$r=\frac{T_{\mathrm{P}}}{T_{\mathrm{P}}+F_{\mathrm{N}}}=\frac{T_{\mathrm{P}}}{P},\tag{13}$$

式中：F_{N} 表示假负例；P 表示图像实际包含的目标总数。召回率反映检测算法能够检测出的目标的比例。IoU 的计算公式为

$$\mathrm{IoU}=\frac{\mathrm{Area}(\mathrm{Box}_y\cap\mathrm{Box}_{\mathrm{gt}})}{\mathrm{Area}(\mathrm{Box}_y\cup\mathrm{Box}_{\mathrm{gt}})},\tag{14}$$

式中：$\mathrm{Area}(\mathrm{Box}_y\cap\mathrm{Box}_{\mathrm{gt}})$ 指的是检测出的目标与真值的平面空间相交的面积；$\mathrm{Area}(\mathrm{Box}_y\cup\mathrm{Box}_{\mathrm{gt}})$ 指的

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

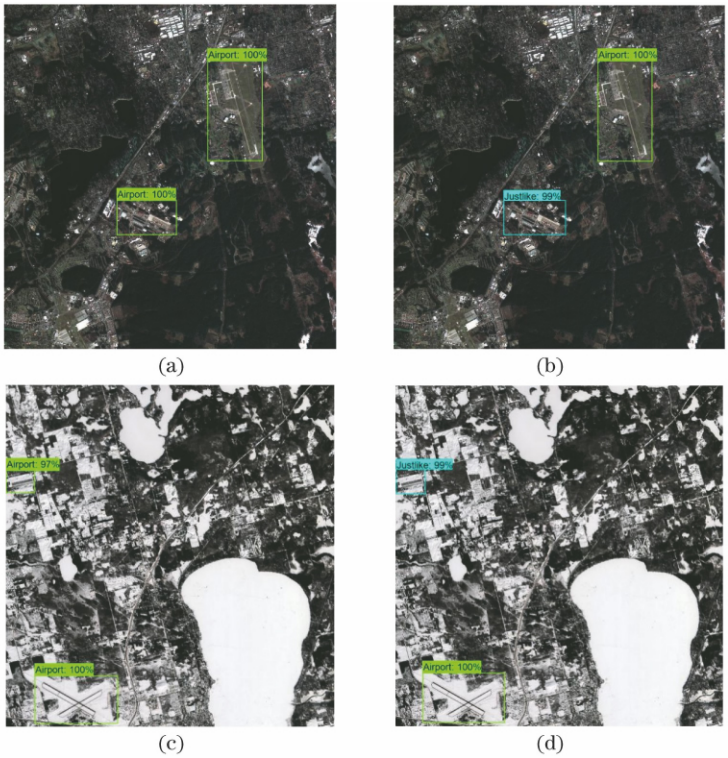

图 10　连续学习模式下机场识别结果的对比。(a) 目标纹理形状与机场相似情况下 CLRNet 1st stage 的识别结果；(b) 目标纹理形状与机场相似情况下 CLRNet 2nd stage 的识别结果；(c) 背景环境干扰下 CLRNet 1st stage 的识别结果；(d) 背景环境干扰下 CLRNet 2nd stage 的识别结果

Fig. 10 Comparison of airport detection results in continuous learning mode. (a) CLRNet 1st stage detection results under the condition that the target texture shape is similar to the airport; (b) CLRNet 2nd stage detection result under the condition that the target texture shape is similar to the airport; (c) CLRNet 1st stage detection result under background environment interference; (d) CLRNet 2nd stage detection result under background environment interference

是检测出的目标与真值的平面空间的并集面积。IoU 指标能很好地反映检测边框定位的准确性。

平均检测精度(mAP^{50})：对测试集检测结果进行统计，如果预测出来的目标与真实框的 IoU 大于阈值 0.5 并且置信度大于 0.8，则认为正确地预测了目标。预测正确的目标数与总预测目标数的比值即为 mAP^{50}。

为了验证本文方法的优越性，选取目前比较流行的几种目标检测方法进行纵向比较，比较结果如表 3 所示。其中：SSD[16] 与 YOLOv3[17] 为基于回归学习的单策略检测模型；Faster R-CNN[18] 为基于区域建议的双策略检测模型；文献[19]所提方法是融合多层次特征的针对遥感目标识别的检测方法。从表 3 中可以发现：连续学习 CLRNet 2nd stage 的平均检测精度相比双策略模型 Faster R-CNN、文献[19]方法分别提升了 3.16%、1.71%，效率提升了 69.3%和 32.3%；单策略模型 SSD 与 YOLOv3 的检测效率比 CLRNet 方法更高，但精度却下降了。

表 3　针对机场目标不同检测方法的精度与效率

Table 3　Precision and efficiency of different detection methods for airport object

Method	Precision	Recall	mAP^{50}	Average detection time/s
SSD[16]	0.9201	0.9355	0.8817	0.12
YOLOv3[17]	0.9322	0.9412	0.8446	0.09
Faster R-CNN[18]	0.9500	0.9631	0.9318	0.75
Method of Ref. [19]	0.9671	0.9465	0.9451	0.34
CLRNet 1st stage	0.9720	0.9581	0.9477	0.23
CLRNet 2nd stage	0.9872	0.9913	0.9613	0.23

基于深度学习的典型目标识别技术可以发挥出机器识别自动化、智能化的优势。在海量遥感时空数据的目标检索、时空分析、选择规划等应用方向上，遥感目标快速识别算法可提供关注目标的地理位置信息。“人工智能”思维正在潜移默化地影响遥感分析，将检测算法与海量遥感数据相结合，可为未来海量遥感数据从数据到信息奠定基础。

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

4 结　论

本文提出了一种以 ResNet 101 为骨干特征网络的连续学习目标检测模型框架 CLRNet。在特征提取网络中，通过增加网络层数使纹理简单、环境干扰小的样本通过权重进行抑制。对于某些机场样本特征比较明确且没有遮挡的正样本，对网络损失的贡献进行抑制；对于某些与目标样本纹理结构比较相似且不易区分的负样本，则对网络损失的贡献进行加权。这样可以使得损失函数的参数梯度诱导模型向复杂的目标进行参数学习，有效提升了关键目标的检测精度。

结合机场目标在遥感影像上的形状与尺度特点，减少锚点缩放尺度，使候选框数量减少，大幅提升了检测效率。在连续学习的基础上构建核心集进行训练，可使模型的泛化能力更强。实验结果表明，本文所提方法对机场目标的平均检测精度比目前流行的基于回归学习的单策略检测方法 YOLOv3、SSD 分别提升了 13.8%与 9.0%，比双策略检测方法 Faster R-CNN 提升了 3.16%。

遥感影像中存在不同尺寸的目标，而不同的目标具有不同的尺度特征。未来拟借鉴无锚点逐像素回归策略检测模型[20]，利用特征金字塔结构的单策略网络进行多尺度识别，同时考虑结合高光谱影像光谱信息进行网络架构设计，以识别更多尺度差异性大且与光谱敏感的遥感目标。将检测结果与实际应用结合起来，可为全球化大数据分析提供支撑。

参　考　文　献

[1] Zhong X, An Y, Wang D, et al. Jilin-1 construction of commercial aerospace remote sensing service system[J]. Satellite Application, 2020(3): 8-17.
钟兴, 安源, 王栋, 等. 吉林一号商业航天遥感服务体系建设[J]. 卫星应用, 2020(3): 8-17.

[2] Wang X, Wang B, Zhang L M. Airport detection in remote sensing images based on visual attention[M]// Neural Information Processing. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 475-484.

[3] Koc-San D, Selim S, Aslan N, et al. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform[J]. Computers and Electronics in Agriculture, 2018, 150: 289-301.

[4] Zhu D, Wang B, Zhang L M. Airport target detection in remote sensing images: a new method based on two-way saliency[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(5): 1096-1100.

[5] Aytekin Ö, Zöngür U, Halici U. Texture-based airport runway detection[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(3): 471-475.

[6] Yu D H, Zhang N, Zhang B M, et al. Airport detection using convolutional neural network and salient feature[J]. Bulletin of Surveying and Mapping, 2019(7): 44-49.
余东行, 张宁, 张保明, 等. 结合卷积神经网络与显著性特征的机场检测[J]. 测绘通报, 2019(7): 44-49.

[7] Liu Y, Yao J, Feng C. An efficient method for airplane detection in high-resolution remote sensing images[J]. Journal of Geomatics, 2020, 45(1): 95-100.
刘媛, 姚剑, 冯辰. 一种高效的高分辨率遥感影像飞机目标检测方法[J]. 测绘地理信息, 2020, 45(1): 95-100.

[8] 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.

[9] 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.

[10] Cai B W, Jiang Z G, Zhang H P, et al. Airport detection using end-to-end convolutional neural network with hard example mining[J]. Remote Sensing, 2017, 9(11): 1198.

[11] Zhu M M, Xu Y L, Ma S P, et al. Airport detection method with improved region-based convolutional neural network[J]. Acta Optica Sinica, 2018, 38(7): 0728001.
朱明明, 许悦雷, 马时平, 等. 改进区域卷积神经网络的机场检测方法[J]. 光学学报, 2018, 38(7): 0728001.

[12] Zhang Z X, Yang C L, Zhu R F, et al. An algorithm for recognition of airport in remote sensing image based on DCNN model[J]. Electronics Optics & Control, 2018, 25(6): 83-89.
张作省, 杨程亮, 朱瑞飞, 等. 联合深度卷积神经网络的遥感影像机场识别算法[J]. 电光与控制, 2018, 25(6): 83-89.

[13] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(13): 3521-3526.

[14] Lee S W, Kim J H, Jun J, et al. Overcoming catastrophic forgetting by incremental moment

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

matching[Z/OL]. 2017: arXiv: 1703. 08475. https://arxiv.org/abs/1703.08475.

[15] Feldman D, Fiat A, Sharir M. Coresets and their applications[D]. Tel-Aviv: Tel-Aviv University, 2010: 100-102.

[16] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]// Computer Vision-ECCV 2016. [S. l.]: Springer, 2016: 21-37.

[17] Redmon J, Farhadi A. YOLOv3: an incremental improvement[Z/OL]. 2018: arXiv: 1804. 02767. https://arxiv.org/abs/1804.02767.

[18] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. [S. l.: s. n.], 2015: 91-99.

[19] Wang X L, Ban Y, Guo H M, et al. Deep learning model for target detection in remote sensing images fusing multilevel features[C]// 2019 IEEE International Geoscience and Remote Sensing Symposium. 28 July-2 Aug. 2019, Yokohama, Japan. New York: IEEE Press, 2019: 250-253.

[20] Tian Z, Shen C H, Chen H, et al. FCOS: fully convolutional one-stage object detection[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 27 Oct.-2 Nov. 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 9626-9635.

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net