

文章编号: 1001-9081(2017)S2-0085-04

# 基于 Faster RCNN 以及多部件结合的机场场面静态飞机检测

戴陈卡<sup>1,2\*</sup> 李毅<sup>1</sup>

(1. 四川大学 计算机学院, 成都 610065; 2. 四川大学 国家空管自动化系统技术重点实验室, 成都 610065)

(\* 通信作者电子邮箱 2471640281@qq.com)

**摘要:** 针对机场场面飞机检测中, 由于目标几何轮廓不完整、姿态不一等原因导致的目标检测难度大、准确率低的问题, 提出了基于 Faster RCNN 以及多部件结合的机场场面静态飞机检测方法。首先, 使用 Faster RCNN 框架提取飞机及其部件(包括尾翼、机翼、机身), 先降低检测置信度, 检测出足够多的部件; 然后, 再根据飞机与部件的交叠率, 判断部件与所属飞机的关系, 滤去孤立的飞机目标和部件; 最后, 调整检测框大小, 获得最终的飞机检测框。实验结果表明, 所提方法相较于原始的 Faster RCNN, 对机场场面的飞机识别率提高了 9%, 对不同姿态、不同场景下的飞机目标检测皆有不错的检测效果, 目标识别准确率达 90% 以上。

**关键词:** 卷积神经网络; 飞机检测; 多部件结合; 目标识别

**中图分类号:** TP183 **文献标志码:** A

## Aeroplane detection in static aerodrome based on Faster RCNN and multi-part model

DAI Chenka<sup>1,2\*</sup>, LI Yi<sup>1</sup>

(1. College of Computer Science, Sichuan University, Chengdu Sichuan 610065, China;

2. National Key Laboratory of Air Traffic Control Automation System Technology, Sichuan University, Chengdu Sichuan 610065, China)

**Abstract:** Concerning the problem that aeroplane detection is difficult and the low accuracy in static aerodrome because of the incomplete geometric profile of the target aeroplane, an aeroplane detection in static aerodrome based on Faster Regions with Convolutional Neural Network (RCNN) and multi-part model was proposed. Firstly, the Faster RCNN was used to extract the aeroplane and its parts (including the tail, wing, body), reduced the detection confidence to get enough parts. And then according to the overlap ratio of the aeroplane and the parts, the relationship between the aeroplane and its parts was determined, and the isolated aeroplane targets and parts were filtered out. Finally the size of the detection box was adjusted to get the final aeroplane detection box. The experimental results show that the proposed method can improve the recognition rate of aeroplane in the aerodrome by 9%, and it has a good detection effect on the aircraft targets under different postures and different scenarios with the target recognition accuracy of over 90%.

**Key words:** Convolutional Neural Network (CNN); aeroplane detection; multi-part model; target recognition

## 0 引言

目标识别是计算机视觉领域的核心问题之一, 而飞机目标的识别又是它在军事领域和民用领域里的一项重要应用。

另一方面, 随着卷积神经网络 (Convolutional Neural Network, CNN) 相关技术的迅速发展, 越来越多的人将 CNN 应用于飞机目标的检测, 比如姚相坤等<sup>[1]</sup>提出基于多结构卷积神经网络的高分遥感影像飞机目标检测方法, 构建了一种多结构卷积神经网络模型去检测遥感图像中的飞机。

机场场面的飞机检测也有很多研究, 比如李圣等<sup>[2]</sup>提出了一种复杂条件下机场飞机的检测方法, 使用方向梯度直方图 (Histogram of Oriented Gradients, HOG) 特征来进行飞机的检测<sup>[3]</sup>, 取得了不错的检测效果。但该方法在处理目标遮挡情况以及运算速度上, 还有可提升的空间。蒋俊提出了一种使用零空间核 Fisher 判别分析 (Kernel Fisher Discriminant Analysis, KFDA)<sup>[4]</sup>方法来进行飞机检测<sup>[5]</sup>。该方法在机型识别上取得了不错的效果, 但仍存在准确率不够高的问题。

针对机场场面飞机检测中, 由于目标几何轮廓不完整、姿态不一等原因导致的目标检测难度大、准确率低的问题, 本文

提出了一种基于卷积神经网络, 使用多部件模型结合<sup>[6-7]</sup>的机场场面飞机检测方法。该方法使用 Faster Regions with Convolutional Neural Network (RCNN) 进行目标的初步检测, 提取飞机以及部件的特征, 再利用多部件结合的方法, 准确地检测出飞机目标。

## 1 Faster RCNN 目标提取

机场场面飞机检测与遥感图像飞机检测的区别在于, 场面的监控图像并不能像遥感航拍图像那样能获取完整的飞机轮廓。视频监控中的飞机, 在大小、姿态上都会有很多的变化, 而且由于监控设备性能的限制, 目标的清晰度也会比较低, 这就大大增加了目标检测的难度, 所以需要一种稳健的方法去实现飞机的检测。本文使用的卷积神经网络为 Faster RCNN<sup>[8]</sup>, 主要思路为区域建议网络 (Region Proposal Network, RPN) + Fast RCNN<sup>[9]</sup>。

### 1.1 特征提取

在图像特征的提取上, 卷积神经网络有着极其优越的特性, 它的权值共享网络结构使之更类似于生物神经网络, 降低了网络模型的复杂度, 减少了权值的数量。另外, 在处理图像

收稿日期: 2017-06-23; 修回日期: 2017-09-04。 基金项目: 十二五国家空管委空管科研课题 (GKG201403004)。

作者简介: 戴陈卡 (1992—), 男, 浙江温州人, 硕士研究生, CCF 会员, 主要研究方向: 深度学习、图像识别; 李毅 (1967—), 男, 四川成都人, 副教授, 博士, 主要研究方向: 空管自动化系统。

时,图像可以直接作为网络的输入,避免了传统识别算法中复杂的特征提取和数据重建过程。

在 Faster RCNN 的框架下,本文使用 VGG16 网络<sup>[10]</sup>进行特征提取。VGG16 网络结构相较于 LeNet<sup>[11]</sup>、AlexNet<sup>[12]</sup>、ZFNet<sup>[13]</sup> 这些常用的卷积神经网络,其特点就是深度更深,能更好地提取特征,取得更好的检测结果。

### 1.2 区域建议网络

区域建议网络(Region Proposal Network, RPN)是一个全卷积网络(Fully Convolutional Network, FCN)<sup>[14]</sup>,用于生成高质量的区域建议框。它和检测网络共享全图的卷积特征,解决了原来的选择性搜索(Selective Search, SS)<sup>[15]</sup>方法的速度瓶颈,大大提高了目标检测速度。

区域建议网络的主要结构如图1所示。

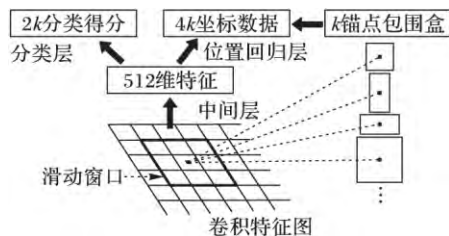


图1 区域建议网络结构

用一个滑动窗口(3×3的卷积核)在一个卷积特征图(feature map)上进行滑动,该特征图由前文介绍的VGG16网络的前5部分卷积层(不包括maxpool层以及之后的全连接层)得到。滑动(卷积)操作获得一个512维的向量后,将这个向量送入到两个平行的全连接层:分类层(box-classification layer (cls))和位置回归层(box-regression layer (reg))以得到分类信息和位置信息。

每个滑动窗口的中心,对应 $k$ 个锚点(anchor)。每个锚点对应一种尺寸和长宽比(如图1右侧所示),RPN使用3种尺寸和长宽比,所以每一个滑动窗口有 $k=9$ 个锚点。相应地,在每一滑动窗口的位置,同时预测9个区域建议,所以位置回归层有4×9个输出,代表9个区域建议框的坐标编码;分类层输出2×9个输出,内容为对每个建议框是目标/非目标的概率。

### 1.3 目标识别

使用区域建议网络生成建议区域之后,Faster RCNN使用Fast RCNN进行检测分类。Fast RCNN是Ross在2015年提出的,该方法解决了RCNN检测速度慢和训练时间空间消耗大的问题,并实现了部分端到端(end-end)的联合训练(joint training),有很好的检测效果。

区域建议网络和Fast RCNN共享之前VGG16前五部分提取的卷积特征,Faster RCNN使用RPN提供的高质量建议区域进行目标识别,大幅提高了目标检测的速度。其网络结构如图2所示。

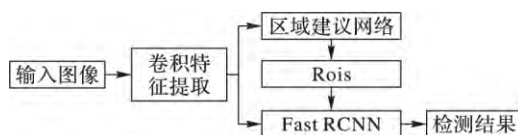


图2 Faster RCNN 结构

由于使用了RPN和Fast RCNN两种网络,不方便直接使用反向传播(Back Propagation, BP)算法进行训练,在训练时,采取轮流训练的方法。

第一阶段,训练RPN。该网络用ImageNet预训练模型

(称之为 $M_0$ )初始化,然后用训练数据进行训练。RPN训练完成后,得到模型 $M_1$ 。接下来使用上一步训练得到的模型 $M_1$ 生成建议区域 $P_1$ 。

第二阶段,使用第一步的RPN生成的建议区域 $P_1$ 来训练Fast RCNN得到模型 $M_2$ ,Fast RCNN模型同样由ImageNet模型 $M_0$ 初始化。

第三阶段,用 $M_2$ 初始化RPN训练得到模型 $M_3$ ,与第一阶段不同的是,这一阶段的训练,固定了共享的卷积层的参数,只是微调了RPN的参数。同样地,使用训练完成的RPN生成建议区域 $P_2$ 。

第四阶段,使用 $M_3$ 以及 $P_2$ 训练Fast RCNN得到最终模型 $M_4$ 。这一阶段的训练,保持共享的卷积层参数固定,只是训练了Fast RCNN的全连接层。这样,两个网络共享相同的卷积层,构成一个统一的网络。具体每个阶段的迭代次数如表1所示。

表1 Faster RCNN 训练过程

训练阶段	网络	迭代次数
1	RPN	40 000
1	Fast RCNN	20 000
2	RPN	40 000
2	Fast RCNN	20 000

## 2 多部件结合

一般飞机可分为机头、机翼、机身和机翼这几个部分,分别对飞机图像中的这几个部件进行标定,然后使用Faster RCNN进行训练,训练结果如表2所示。

表2 飞机各部分检测准确率

部件	平均准确率
飞机	0.9057
机翼	0.8044
尾翼	0.9056
机身	0.8000

由表2可知:在被检测图片中,飞机本身以及尾翼平均准确率(Average Precision, AP)最高,最容易被检测出,而机体以及机翼的被识别率相对较低。在不同的条件下(视角、是否有遮挡、光照),飞机各部件会在目标检测时有不同的贡献。如果一个区域被检测为飞机,而这个区域中又有飞机部件被检测出来,那么这个区域中的目标为飞机的概率就大大增加。为了利用这一特性,本文设计使用了一种多部件结合的飞机目标检测方法,其算法流程如图3所示。

算法开始之后,使用Faster RCNN对输入图像进行目标检测。为了在初检阶段获得更多的飞机以及部件目标,算法在初检时使用了一个较低的置信度阈值,使得初步检测时,更多的目标可以被检测出来,比如图4(a)(其中一个飞机检测框的置信度为0.666);但相对地,也有很多错误的目标被标

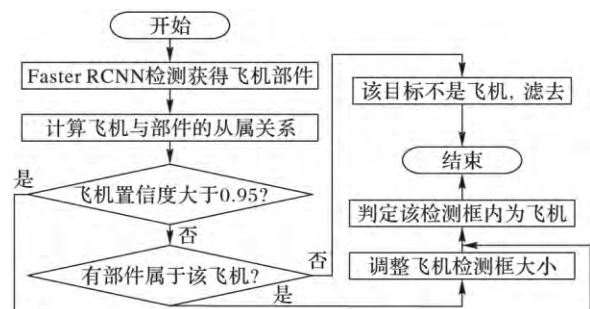


图3 多部件结合检测算法流程

算法开始之后,使用Faster RCNN对输入图像进行目标检测。为了在初检阶段获得更多的飞机以及部件目标,算法在初检时使用了一个较低的置信度阈值,使得初步检测时,更多的目标可以被检测出来,比如图4(a)(其中一个飞机检测框的置信度为0.666);但相对地,也有很多错误的目标被标

记出来,如图 4(b),一个廊桥被误检为飞机。这时就需要对检测出的目标进行筛选。



(a) 检出低置信度目标

(b) 存在目标误检

图 4 初始检测结果

这里先对初检的置信度阈值  $\omega$  进行讨论。当检测框的置信度大于  $\omega$  时,代表该检测框内的物体为网络所检测的目标;当置信度小于  $\omega$  时,则代表检测框内没有检测目标。经过实验,不同置信度阈值下,飞机目标的漏检率与错检率如表 3 所示。

表 3 不同置信度下目标检测的准确率

置信度阈值	漏检率/%	错检率/%
0.4	13.0	6.4
0.3	11.3	10.1
0.2	8.7	17.4
0.1	8.7	23.1

由表 3 可以看出,当置信度阈值为 0.2 时,漏检率最小,虽然错检率随着置信度要求的下降而升高,但错检的目标可以使用多部件结合的方法去除,所以本文使用了 0.2 作为初检置信度阈值,以求最大限度地提升检测的准确率。

在初检之后,本文使用多部件结合的方法,去除错误的检测结果。首先将初检获得的检测框分为两类:飞机与飞机部件。然后计算每个飞机与所有部件之间的交叠率,确定部件与飞机的从属关系。部件属于与其  $IoU$  最大的飞机。 $IoU$  的计算公式如下:

$$IoU = \frac{PartAcreage \cap PlaneAcreage}{PartAcreage \cup PlaneAcreage}$$



图 5 多部件结合

完成飞机与部件的配对之后,滤去孤立的部件模型以及飞机,剩下的就是希望检测到的飞机目标,如图 5 所示。图 5 是图 4(b) 经过多部件处理之后的图,被误检为飞机的廊桥,由于没有部件与其相交 ( $IoU = 0$ ),被判断为孤立的飞机,最终被滤去。剩下的两架飞机,都完成了部件的匹配(见图 5 连线)。

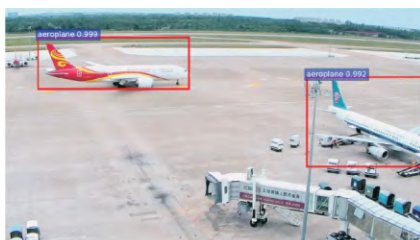


图 6 最终检测结果

上一步得到的检测结果就是要检测的飞机目标,最后根

据飞机所属部件的位置,微调飞机检测框的大小(尽量使飞机检测框与部件检测框外接),然后去掉部件的检测框,得到最后的检测结果。如图 6 所示,在最终的检测结果中,多部件结合方法,成功地筛选了错检的目标,保留了正确的检测结果。

### 3 实验结果与分析

本文使用的仿真平台为 Linux,深度学习框架为 Caffe,共使用了 1 000 张图片进行训练,先将图片归一化到宽度为 800,长度 500 ~ 600(单位:像素)的大小,之后在尺寸归一化的图片上进行飞机各部件的人工标定(如图 7 所示)。标定完成后使用 Faster RCNN(使用 VGG16 网络提取卷积特征)进行训练,迭代次数如表 1 所示,学习率为 0.001。



图 7 多部件模型

如图 8(a) 所示,使用原始的 Faster RCNN,由于角度以及图像像素的问题,有一架飞机没有被检测出来。而使用了多部件结合的方法,由于该飞机的尾翼特征较为明显(置信度 0.717),确定了部件与飞机的关系,成功检测出了飞机目标(如图 8(b) 所示)。

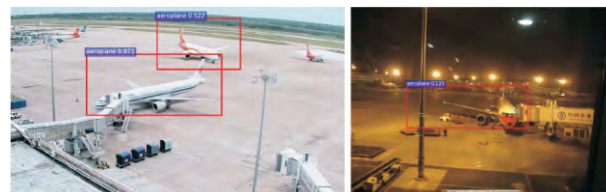


(a) 原始方法

(b) 多部件结合

图 8 原始方法与多部件结合方法的对比

部分检测结果如图 9 所示,本文使用的方法,对白天、夜间不同姿态的飞机目标,都有较好的检测结果。



(a) 白天的检测结果

(b) 夜间的检测结果

图 9 不同光照条件下的检测结果

得到训练模型之后,本文使用了 120 张图片进行测试,测试图片来源主要为成都双流机场的场面监控视频。除了原始 Faster RCNN 方法,本文与 KFPA 方法和 HOG 特征方法进行了对比实验,实验结果显示:KFPA 方法的检测准确率为 78.3%,HOG 特征方法的检测准确率为 81.8%,原始 Faster RCNN 的检测准确率为 83.4%,本文方法的检测准确率为 91.3%。

由实验结果可以得出,本文方法在机场场面静态飞机的识别率上优于对比算法。同时可以分析得到,在机场场面监控获得的图像中,由于场景的复杂度变大、飞机在场景中的比例大幅减小,以及监控视频的清晰度问题,使得飞机目标检测

的难度大大增加,准确率相较于单一场景的目标检测有了一定的程度的下降(对比表2中训练数据90.57%的准确率),而本文提出的多部件结合的目标检测方法,在一定程度上解决了这个问题,对不同角度、不同光照条件下的飞机目标有较好的检测效果(如图9所示),提升了目标检测的准确率。

#### 4 结语

针对机场场面飞机检测中,由于目标几何轮廓不完整、姿态不一等特性,导致的目标检测难度大、准确率低的问题,提出了基于Faster RCNN以及多部件结合的机场场面静态飞机检测方法。与其他方法的对比实验表明,本文使用的方法,提升了大场景下飞机检测的准确率,对复杂场景下不同姿态的飞机检测有较好的效果。但是,本文方法对场景中远距离小飞机的检测效果并不好(如图9(a)远处中的小飞机),导致本文方法没有取得更高的准确率。由于机场监视设备的限制,远处的飞机往往较为模糊,难以提取到部件特征,无法进行优化。所以,在以后的研究中,将重点解决机场场面中远处目标的检测问题,包括图像特征的加强、小目标位置的回归等,以获得更好的检测效果。

#### 参考文献:

- [1] 姚相坤,万里红,霍宏,等. 基于多结构卷积神经网络的高分遥感影像飞机目标检测[J]. 计算机工程,2017,43(1):259-267.
- [2] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC: IEEE Computer Society, 2005, 1: 886-893.
- [3] 李圣,郑宏,张博凯,等. 一种复杂条件下机场飞机的检测方法[J]. 科学技术与工程,2015,15(23):43-49.
- [4] YANG J, FRANGI A F, YANG J, et al. KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(2): 230-244.
- [5] 蒋俊. 白云机场目视泊位系统飞机识别的改进方法与研究[J]. 信息通信,2016(7):74-75.
- [6] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]// CVPR 2008: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2008: 1-8.
- [7] ZHANG Y, HE C, SHI B, et al. Multi-layer feature selection based hierarchical component model for aeroplane detection on remote sensing image[J]. Geomatics & Information Science of Wuhan University, 2014, 39(12):1406-1411.
- [8] REN S, HE K, 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.
- [9] GIRSHICK R. Fast R-CNN[C]// Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1440-1448.
- [10] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint, 2014, 2014: arXiv:1409.1556.
- [11] HAYKIN S, KOSKO B. Gradient based learning applied to document recognition[M]. Hoboken: Wiley-IEEE Press, 2009: 306-351.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems. Redhook: Curran Associates Inc., 2012: 1097-1105.
- [13] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]// Proceedings of the 2014 European conference on Computer Vision, LNCS 8689. Cham: Springer International Publishing AG, 2014: 818-833.
- [14] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.
- [15] UIJLINGS J R R, VAN DE SANDE K E A, GEVERS T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154-171.

(上接第67页)

#### 参考文献:

- [1] WANG X, JIANG Z, YU D. An improved KNN algorithm based on kernel methods and attribute reduction[C]// IMCCC 2015: Proceedings of the 5th International Conference on Instrumentation and Measurement, Computer, Communication, and Control. Piscataway: IEEE, 2015: 567-570.
- [2] ZHANG S C, LI X L, ZONG M, et al. Learning k for kNN classification[J]. ACM Transactions on Intelligent Systems and Technology, 2017, 8(3): 43.
- [3] 李秀娟. KNN分类算法研究[J]. 科技信息, 2009(31): 81.
- [4] 王增民,王开珏. 基于熵权的K最近邻算法改进[J]. 计算机工程与应用,2015,45(30):129-131.
- [5] 孙可,龚永红,邓振云. 一种高效的K值自适应的SA-KNN算法[J]. 计算机工程与科学,2015,37(10):1965-1970.
- [6] 余小鹏,周德翼. 一种自适应k-最近邻算法的研究[J]. 计算机应用研究,2006,23(2):70-72.
- [7] DUDAI S A. The distance-weighted k-nearest neighbor rule[J]. IEEE Transactions on Systems, Man and Cybernetics, 1976, 6(4): 325-327.
- [8] 杨金福,宋敏,李明爱. 一种新的基于距离加权的模板约简K近邻算法[J]. 电子与信息学报,2011,33(10):2378-2383.
- [9] 茅剑,刘晋明,曹勇. 一种基于密度的改进KNN文本分类算法[J]. 漳州师范学院学报(自然科学版),2012(2):45-48.
- [10] LI C Q, JIANG L X, LI H, et al. Attribute weighted value difference metric[C]// ICTAI 2013: Proceedings of the 25th International Conference on Tools with Artificial Intelligence. Washington, DC: IEEE Computer Society, 2014: 575-580.
- [11] 肖辉辉,段艳明. 基于属性值相关距离的KNN算法的改进研究[J]. 计算机科学,2013,40(11A):157-159.
- [12] 周志华. 机器学习[M]. 北京:清华大学出版社,2016:101-103.