

基于深度卷积神经网络的遥感图像飞机目标检测方法

郭智^{①②} 宋萍^{*①②③} 张义^{①②} 闫梦龙^{①②} 孙显^{①②} 孙皓^{①②}

^①(中国科学院电子学研究所 北京 100190)

^②(中国科学院空间信息处理与应用系统技术重点实验室 北京 100190)

^③(中国科学院大学 北京 100049)

摘要: 飞机检测是遥感图像分析领域的研究热点, 现有检测方法的检测流程分为多步, 难以进行整体优化, 并且对于飞机密集区域或背景复杂区域的检测精度较低。针对以上问题, 该文提出一种端到端的检测方法 MDSSD 来提高检测精度。该方法基于单一网络目标多尺度检测框架(SSD), 以一个密集连接卷积网络(DenseNet)作为基础网络提取特征, 后面连接一个由多个卷积层构成的子网络对目标进行检测和定位。该方法融合了多层次特征信息, 同时设计了一系列不同长宽比的候选框, 以实现不同尺度飞机的检测。该文的检测方法完全摒弃了候选框提取阶段, 将所有检测流程整合在一个网络中, 更加简洁有效。实验结果表明, 在多种复杂场景的遥感图像中, 该方法能够达到较高的检测精度。

关键词: 遥感图像处理; 飞机检测; 密集连接卷积网络

中图分类号: TP753

文献标识码: A

文章编号: 1009-5896(2018)11-2684-07

DOI: [10.11999/JEIT180117](https://doi.org/10.11999/JEIT180117)

Aircraft Detection Method Based on Deep Convolutional Neural Network for Remote Sensing Images

GUO Zhi^{①②} SONG Ping^{①②③} ZHANG Yi^{①②} YAN Menglong^{①②}
SUN Xian^{①②} SUN Hao^{①②}

^①(*Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China*)

^②(*Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China*)

^③(*University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Aircraft detection is a hot issue in the field of remote sensing image analysis. There exist many problems in current detection methods, such as complex detection procedure, low accuracy in complex background and dense aircraft area. To solve these problems, an end-to-end aircraft detection method named MDSSD is proposed in this paper. Based on Single Shot multibox Detector (SSD), a Densely connected convolutional Network (DenseNet) is used as the base network to extract features for its powerful ability in feature extraction, then an extra sub-network consisting of several feature layers is appended to detect and locate aircrafts. In order to locate aircrafts of various scales more accurately, a series of aspect ratios of default boxes are set to better match aircraft shapes and combine predictions deduced from feature maps of different layers. The method is more brief and efficient than methods that require object proposals, because it eliminates proposal generation completely and encapsulates all computation in a single network. Experiments demonstrate that this approach achieves better performance in many complex scenes.

Key words: Remote sensing image processing; Aircraft detection; Densely connected convolutional network

收稿日期: 2018-01-26; 改回日期: 2018-06-06; 网络出版: 2018-08-30

*通信作者: 宋萍 pingsong2014@163.com

基金项目: 国家自然科学基金(41501485)

Foundation Item: The National Natural Science Foundation of China (41501485)

1 引言

高分辨率遥感图像中飞机目标检测是遥感图像处理领域的研究热点,在军事和民用方面都应用广泛。为了实现遥感图像中的飞机目标检测,研究学者们相继提出了许多有价值的目标检测方法^[1,2]。

目前的飞机检测算法通常可以分为两类:分阶段检测方法和端到端检测方法。分阶段检测方法一般先通过滑动窗等方法获得候选区域,之后提取特征训练分类器,再通过分类器判断候选框中是否包含目标。目前遥感领域许多目标检测算法都是基于这种方法实现的^[3-6]。这类检测方法的检测流程均包含多个阶段,实现过程复杂,并且候选区域提取方法冗余性很大。Faster R-CNN^[7]和SSD^[8]都是基于回归方法的端到端目标检测算法,在计算机视觉领域应用广泛。Faster R-CNN仍包含耗时的候选框提取阶段,因此还达不到实时检测。SSD是基于深度卷积网络的检测算法,将候选框提取阶段和类别预测阶段有效结合,提高了检测速度,但对小目标的检测效果较差。这类检测方法主要针对自然场景设计,而遥感图像中的目标具有尺度小,分辨率低的特点。因此,不能直接适用于遥感图像。

结合遥感图像中飞机目标的特点,本文提出了一种基于密集连接网络的多尺度检测框架(Multiscale Densenet SSD, MDSSD)。本文的主要工作包括以下两点:(1)以SSD的目标检测框架为基础,引入DenseNet^[9]作为基础网络提取特征,DenseNet的密集连接结构(dense block)加强了特征传递,可以更有效地利用特征,缓解了深度卷积网络容易出现的梯度消失现象;(2)根据不同型号飞机特征设计了用于检测目标的特征层大小以及候选框的长宽比,通过融合不同层次特征层的检测结果实现飞机的多尺度检测。

2 本文目标检测方法

2.1 网络结构

本文提出的MDSSD检测模型网络结构如图1所示,网络结构参数见表1。本文的检测模型使用DenseNet作为基础网络提取特征。文献^[10]证明了网络深度是提高分类准确率的关键因素。然而,网络深度的增加容易导致降级问题,即准确率趋于平稳后突然降低。DenseNet通过引入密集连接结构减轻了深度网络的降级问题,同时DenseNet减少了每层网络的计算量,并通过密集连接结构进行特征的重复利用,使得每层网络只需学习很少的特征,显著减少了参数量和计算量,极大地提升了效率^[9]。

MDSSD所使用的基础网络为50层的DenseNet,

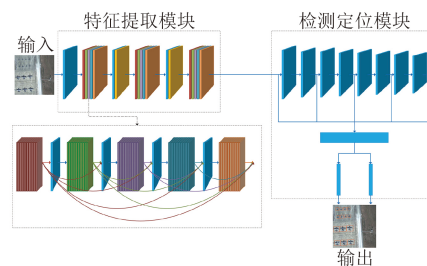


图1 MDSSD检测模型的网络结构示意图

表1 MDSSD检测模型网络参数列表

网络层	网络层参数	输出特征图大小
卷积层	7×7 conv, 步长2	256×256
密集连接1	$[1 \times 1$ conv, 3×3 conv] $\times 9$	256×256
转换层1	1×1 conv	256×256
	2×2 平均池化, 步长2	128×128
密集连接2	$[1 \times 1$ conv, 3×3 conv] $\times 9$	128×128
转换层2	1×1 conv	128×128
	2×2 平均池化, 步长2	64×64
密集连接3	$[1 \times 1$ conv, 3×3 conv] $\times 9$	64×64
卷积层Conv1_1	1×1 conv	64×64
卷积层Conv1_2	3×3 conv, 步长2	32×32
卷积层Conv2_1	1×1 conv	32×32
卷积层Conv2_2	3×3 conv, 步长2	16×16
卷积层Conv3_1	1×1 conv	16×16
卷积层Conv3_2	3×3 conv, 步长2	8×8
卷积层Conv4_1	1×1 conv	8×8
卷积层Conv4_2	3×3 conv, 步长2	4×4

其增长率为12,即每个密集连接结构中每层输出的特征图数目为12^[9],密集连接结构详解图如图1特征提取模块所示。由表1可知,密集连接结构中的每个 3×3 卷积层前都包含一个 1×1 的卷积层,即“瓶颈”层,目的是减少输入特征图的数量,在降维的同时,又可以融合各个通道的特征。每两个密集连接结构之间的转换层中的 1×1 卷积层则是为了进一步压缩网络参数,输出的通道数为输入通道数的一半。转换层中的池化层则使特征图的大小减半。在密集连接结构中每一层的输入来自前面所有层的输出,这种连接方式使得特征和梯度的传递更加有效,网络也更加容易收敛。同时,这种连接方式使得每层都直接连接输入并直接影响损失函数的梯度,因此可以减轻深度卷积网络容易出现的梯度消失现象。

在基础网络后连接了一个包含多个卷积层的子网络用于检测和定位,检测定位模块的网络参数见表1。子网络中每个 3×3 卷积层前面也都包含一个 1×1 的卷积层,其目的仍然是为了降低特征图的

维数, 步长为2的卷积操作则使得经过这些卷积层的特征图大小依次减半。

2.2 多尺度检测

传统方法使用图像金字塔结构实现目标的多尺度检测, 但这种方法的时间复杂度较高。在SSD检测算法的启发下, 本文通过融合网络中不同层次特征层的检测结果来实现不同尺度目标的检测。不同

大小特征图中单位像素的感受野不同, 越小的特征图中单位像素的感受野越大^[1], 因而高层特征图的候选框适用于检测规格较大的飞机, 低层特征图中的候选框适用于检测较小的飞机。以图2为例, 特征图 8×8 适用于检测较小的飞机(图2(b)中黄色矩形框), 特征图 4×4 适用于检测较大的飞机(图2(c)中红色矩形框)。

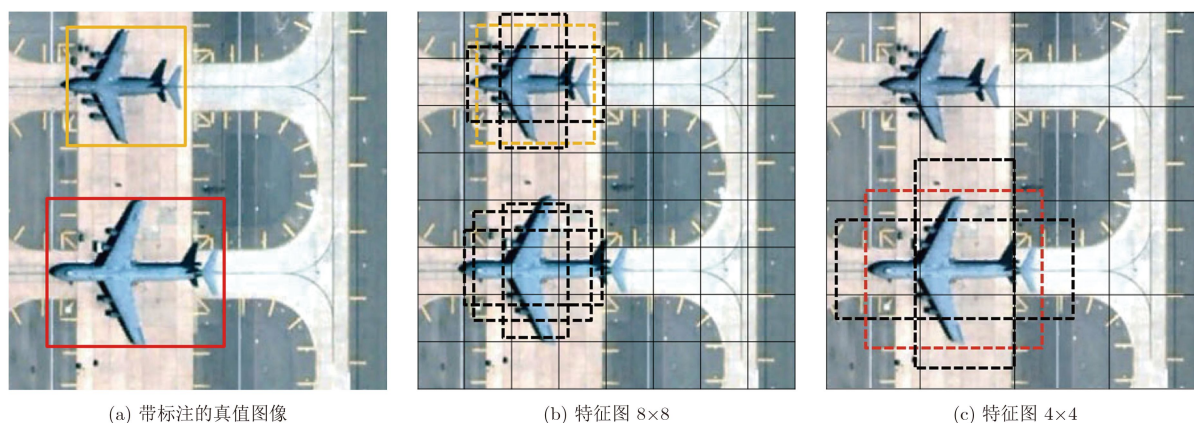


图2 飞机目标多尺度检测示例

常见不同型号飞机的真实尺寸如表2所示。MDSSD是根据表2中飞机的长度和翼展范围选取用于检测的特征层, 即每个来源于不同特征层的候选框都是用于检测某个特定尺寸范围的飞机目标, 用于检测的特征层及大小可从图1和表1获知。

表2 常见飞机型号参数列表

飞机型号	飞机长度 (m)	飞机翼展 (m)	飞机长度/翼展 范围(m)
F-16	15.09	9.45	10~20
MV-22	17.50	14.00	
AH-64	17.76	14.63	
F-22	18.90	13.56	20~40
C-130H	21.00	30.00	
P-3C	35.57	30.36	
KC-135	41.51	39.87	40~60
B-1B	44.50	24.00	
B-52H	49.05	56.40	
C-17	53.29	50.29	60~80
C-5	75.54	67.88	
BOEING747-8	76.40	68.50	
Antonov An 225	84.00	88.40	80~100

2.3 候选框的设计

假设选取 m 个特征层用于检测目标, 则第 k 个特征层的候选框尺度系数为

$$s_k = \begin{cases} s_{\min}, & k = 1 \\ 2s_{\min} + \frac{s_{\max} - 2s_{\min}}{m - 2} \cdot (k - 2), & k \in [2, m] \end{cases} \quad (1)$$

$$z_{\min} = a \cdot s_{\min} \quad (2)$$

$$z_{\max} = a \cdot s_{\max} \quad (3)$$

其中, s_{\min} 和 s_{\max} 分别表示尺度系数的最小值和最大值, z_{\min} 和 z_{\max} 分别表示飞机长度/翼展在图像中所占像素的最小值和最大值, a 表示输入图像的边长。由表2可知, 飞机真实尺寸范围在15~100 m之间(这里对于每种型号的飞机, 取长度和翼展中较大的值作为飞机尺寸), 根据图像分辨率即可获得相应的像素大小, 本文遥感图像数据的分辨率均为0.5 m, 由此可知 z_{\min} 为30像素, z_{\max} 为200像素。由式(2)和式(3)可知, s_{\min} 为0.06, s_{\max} 为0.39。再根据式(1)即可计算出其他几个特征层的尺度系数。

由表2可知, 常见型号飞机的长宽比处于0.5~2.0之间。由此, 本文设计了一系列候选框的长宽比 r , $r \in \left\{1, \frac{2}{3}, \frac{3}{2}, \frac{1}{2}, 2\right\}$ 。每个候选框对应到原图所占的宽为 $w_k = a \cdot s_k \cdot \sqrt{r}$, 高为 $h_k = (a \cdot s_k) / \sqrt{r}$ 。当长宽比为1时增加一个宽为 $w_k = a \cdot \sqrt{s_k \cdot s_{k+1} \cdot r}$, 高为 $h_k = a \cdot \sqrt{(s_k \cdot s_{k+1}) / r}$ 的候选框, 即在特征图的每个位置都将产生6个候选框。每个候选框的中心点为 $((a \cdot (i + 0.5)) / l, (a \cdot (j + 0.5)) / l)$, $i, j \in [0, l)$, 其中 l 表示第 k 个特征图的边长。

2.4 网络模型训练

在训练检测网络时需要输入图像以及图像中标记每架飞机位置的真值框信息。以图2为例, 图2(a)中的图像经过检测网络后, 在 8×8 和 4×4 两个卷积层中每个位置产生3个候选框。对于每个候选框, 要同时计算候选框中心点相对真值框中心点的偏移以及候选框中包含目标的置信度。在训练阶段, 首先将所有候选框与两个真值框按照匹配算法进行匹配, jaccard匹配算法^[12], 将与真值框的匹配系数大于0.5的候选框视为匹配框。图2(b)中黄色矩形框和图2(c)中红色矩形框是与真值框相匹配的候选框, 它们将被标记为正样本, 其他候选框将被标记为负样本, 正样本用 c^1 表示, 负样本用 c^0 表示。

在网络训练过程中, 产生的损失包括分类损失和位置回归损失。损失函数公式为

$$L(p, d, g) = (1/N) (L_{cls}(p) + L_{loc}(d, g)) \quad (4)$$

其中, N 代表正样本的数目, L_{cls} 表示分类损失函数, L_{loc} 表示位置回归损失函数。

分类损失 $L_{cls}(p)$ 是二分类的softmax损失, 分类时属于飞机类别的置信度用 p^1 表示, 属于背景类别的置信度用 p^0 表示, 则分类损失函数的公式为

$$L_{cls}(p) = - \sum_{i \in c^1} \ln(\hat{p}_i^1) - \sum_{i \in c^0} \ln(\hat{p}_i^0) \quad (5)$$

其中, $\hat{p}_i^1 = \frac{\exp(p_i^1)}{\exp(p_i^0) + \exp(p_i^1)}$, $\hat{p}_i^0 = 1 - \hat{p}_i^1$ 。

位置回归损失函数 $L_{loc}(d, g)$ 是匹配候选框(d)与真值框(g)的Smooth L1损失^[13]。参考Faster R-CNN的位置回归算法, 本文计算匹配候选框与真值框中心点坐标以及宽、高的偏移回归为^[7]

$$L_{loc}(d, g) = \sum_{i \in c^1} \sum_{k \in \{c_x, c_y, w, h\}} \text{Smooth L1}(d_i^k - g_j^k) \quad (6)$$

其中, i 表示第 i 个匹配的候选框, j 表示第 j 个真值框。

3 实验与分析

3.1 数据集与训练参数

本文实验中使用的图像数据为高分辨率光学遥感图像, 数据来源于谷歌地图机场区域以及某飞机坟场区域(位于 $32^\circ 9' 2.73'' N$, $110^\circ 49' 27.85'' W$ 附近), 图像分辨率为0.5 m。为便于训练, 所有图像均被裁切为固定尺寸 512×512 像素, 其中包含飞机的图像将被保留下来作为用于网络训练的实验数据。本文的数据以VOC数据格式标注, 图3是几张经过裁切后的原始图像样本示例。

本文将原始数据集通过旋转(90° , 180° , 270°)以及水平镜像翻转的方式进行数据集扩充。经过数据集扩充之后, 图像总数为14856幅, 随机抽取20%作为测试集, 其余图像数据作为训练集。本文模型训练和测试所使用的硬件为NVIDIA TESLA K80 GPUs, 所使用的深度学习框架为CAFFE框架。

实验中, 模型训练的初始学习率为0.01, 优化方式为随机梯度下降, 动量为0.9, 权值衰减为0.0005, 每次处理32幅图像, 最大迭代次数为60000次。前30000次的学习率为0.01, 后30000次的学习率为0.001。

3.2 实验结果与分析

3.2.1 模型主要参量设计对检测性能的影响 探究主要参量设计对检测性能的影响, 所采用的评价指标为平均准确率(Average Precision, AP)^[14]。

(1) 融合的特征层对检测性能的影响: 表3展示了MDSSD检测模型融合不同特征层对检测性能的影响, 从实验结果可以看出融合 64×64 , 32×32 , 16×16 , 8×8 , 4×4 这5个特征层时检测性能达到最佳, 再增加特征层 2×2 并没有带来性能的提升, 这是因为飞机在分辨率为0.5 m, 大小为 512×512 的图像中所占的比例达不到1/4以上。额外融合特征层反而带来候选框的冗余, 因此本文算法中所设计融合的特征层大小是合理的。

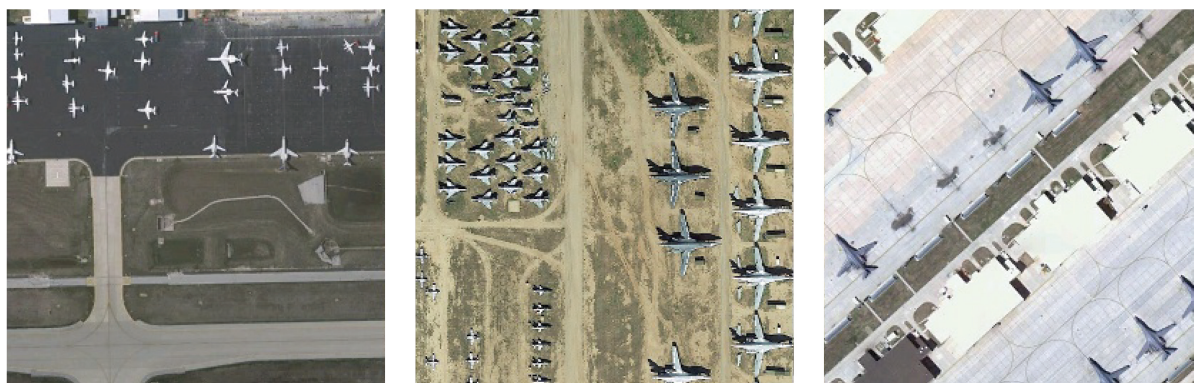


图3 原始图像示例

表3 融合的特征层大小对检测性能的影响

融合的特征层大小						AP (%)
64 × 64	32 × 32	16 × 16	8 × 8	4 × 4	2 × 2	
√	√	√	√	√	√	92.07
√	√	√	√	√		92.07
√	√	√	√			87.42
√	√	√				80.71

(2) 候选框设计对检测性能的影响：由表4的实验结果可以看出设计更多的候选框比例可以取得更好的检测性能。但是增加 $\left\{\frac{1}{3}, 3\right\}$ 这个比例的候选框，并没有提高飞机的检测精度，这是由于遥感图像中飞机的长宽比并不存在这个比例(或者极为少见)，因此增加这个比例的候选框是没有必要的，这证明了本文检测算法中所设计的候选框比例 $\left\{1, \frac{2}{3}, \frac{3}{2}, \frac{1}{2}, 2\right\}$ 的合理性。

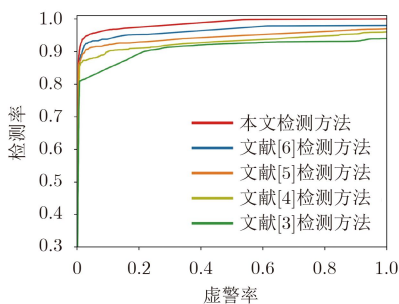
表4 候选框长宽比对检测性能的影响

候选框的长宽比				AP (%)
{1}	$\left\{\frac{1}{2}, 2\right\}$	$\left\{\frac{2}{3}, \frac{3}{2}\right\}$	$\left\{\frac{1}{3}, 3\right\}$	
√	√	√	√	92.07
√	√	√		92.07
√	√			90.47
√				86.53

3.2.2 MDSSD模型与不同检测模型性能对比 为了证明本文提出的目标检测模型MDSSD的有效性，展开了多组对比实验。

(1)对比遥感领域目标检测模型：这里采用检测率和虚警率来评估算法的检测性能，检测率和虚警率的计算公式为

$$\text{检测率} = \frac{\text{检测正确的目标数目}}{\text{测试集中的目标总数}} \quad (7)$$



(a) 各检测模型的 ROC 曲线图

$$\text{虚警率} = \frac{\text{检测错误的目标数目}}{\text{检测出的目标总数}} \quad (8)$$

表5和图4展示了MDSSD与Bag-of-Words模型^[3]，Rotation Invariant Parts-Based模型^[4]，Saliency-Based DBN模型^[5]和ACF+Adaboost模型^[6]的检测性能对比结果。从检测结果可以看出MDSSD在检测率和虚警率两个性能指标上的表现均优于其他方法。Bag-of-Words与Rotation Invariant Parts-Based模型都是利用人工设计的滤波器进行特征提取，更易受到环境等因素的影响，因此相比于Saliency-Based DBN模型的检测效果较差。深度置信网络(DBN)容易过拟合，泛化能力较弱，而ACF+Adaboost模型的检测精度，以及算法的鲁棒性均不够理想。MDSSD能取得良好的检测效果归功于DenseNet的密集连接结构加强了特征的传递，可以更加有效地利用特征，同时有针对性地融合不同层次特征层的检测结果可以提高检测精度。

(2)对比计算机视觉领域目标检测模型：本节将MDSSD与SSD^[8]和Faster R-CNN^[7]进行了对比实验，使用相同的数据集对3个网络进行训练和测试，检测结果如表6和图5所示。表6中所采用的是计算机视觉领域针对目标检测常用的性能评价指标：平均准确率(Average Precision, AP)^[14]。

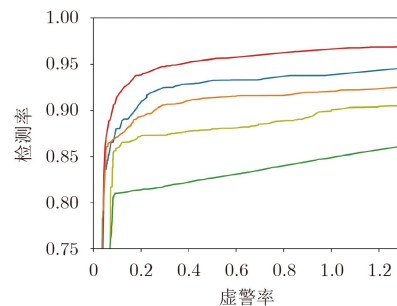
从表6和图5不同模型检测结果可以看出，

表5 MDSSD与遥感领域检测模型性能对比

目标检测模型	虚警率(%)	检测率(%)
Bag-of-Words模型 ^[3]	12.3	84.2
Rotation Invariant Parts-Based模型 ^[4]	11.7	86.2
Saliency-Based DBN模型 ^[5]	9.8	90.1
ACF+Adaboost模型 ^[6]	9.6	92.0
MDSSD	3.2	96.3

表6 3种目标检测模型性能对比

检测模型	SSD	Faster R-CNN	MDSSD
AP(%)	87.64	89.73	92.07



(b) 左图 ROC 曲线局部放大图

图4 五种检测模型的ROC曲线图

MDSSD模型的检测性能优于Faster R-CNN和SSD。由于Faster R-CNN和SSD都是针对自然场景设计的目标检测方法，不能很好地应对遥感图像中飞机尺度小的特点，因此检测结果较差。MDSSD充分结合遥感图像特点，设计了不同层次特征融合方法以及不同候选框尺度，因此对于遥感图像中的飞机目标检测更加有效。

3.2.3 检测结果展示 图6展示了MDSSD模型在不同场景下的检测结果示例。由检测结果可以看出，本文的检测模型在小飞机密集停靠的区域(如图6(a)，图6(c)，图6(d))、或是背景复杂的区域(如图6(b))，亦或是存在残缺飞机并且飞机形状各异的飞机坟场区域(如图6(e)，图6(f))，都可以精确地检测出其中的飞机目标，实验结果证明了本文检测方法的有效性。

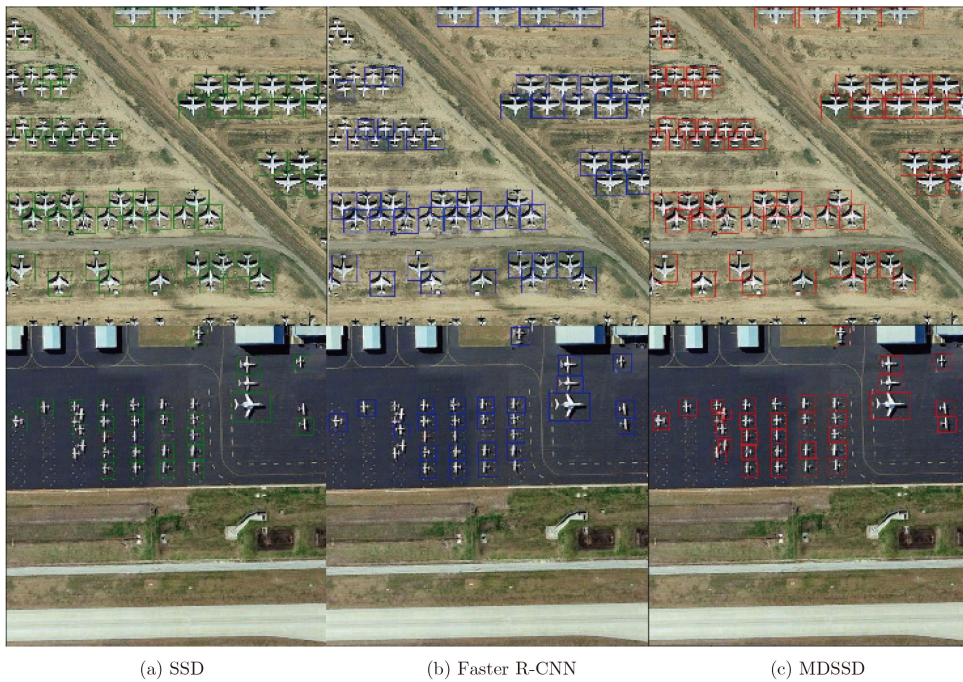


图 5 不同模型检测结果对比图



图 6 本文检测模型的检测结果示例

4 结束语

本文提出了一种基于深度卷积神经网络的端到端目标检测模型MDSSD, 应用于高分辨率遥感图像中的飞机目标检测。该模型以DenseNet作为基础网络提取特征, 其后连接多个卷积层用于检测和定位。本方法根据飞机的真实尺寸设计了用于检测的特征层大小以及候选框的长宽比, 使得模型可以很好地实现多种尺度目标的检测。实验结果表明, 本文的目标检测模型可以准确地检测出遥感图像中的飞机目标, 对比实验证明了本文方法的有效性。

参 考 文 献

- [1] 冯卫东, 孙显, 王宏琦. 基于空间语义模型的高分辨率遥感图像目标检测方法[J]. 电子与信息学报, 2013, 35(10): 2518–2523. doi: [10.3724/SP.J.1146.2013.00033](https://doi.org/10.3724/SP.J.1146.2013.00033).
FENG Weidong, SUN Xian, and WANG Hongqi. Spatial semantic model based geo-objects detection method for high resolution remote sensing images[J]. *Journal of Electronics & Information Technology*, 2013, 35(10): 2518–2523. doi: [10.3724/SP.J.1146.2013.00033](https://doi.org/10.3724/SP.J.1146.2013.00033).
 - [2] 王思雨, 高鑫, 孙皓, 等. 基于卷积神经网络的高分辨率SAR图像飞机目标检测方法[J]. 雷达学报, 2017, 6(2): 195–203. doi: [10.12000/JR17009](https://doi.org/10.12000/JR17009).
WANG Siyu, GAO Xin, SUN Hao, *et al.* An aircraft detection method based on convolutional neural networks in high-resolution SAR images[J]. *Journal of Radars*, 2017, 6(2): 195–203. doi: [10.12000/JR17009](https://doi.org/10.12000/JR17009).
 - [3] SUN Hao, SUN Xian, WANG Hongqi, *et al.* Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model[J]. *IEEE Geoscience and Remote Sensing Letters*, 2012, 9(1): 109–113. doi: [10.1109/LGRS.2011.2161569](https://doi.org/10.1109/LGRS.2011.2161569).
 - [4] ZHANG Wancang, SUN Xian, FU Kun, *et al.* Object detection in high-resolution remote sensing images using rotation invariant parts based model[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(1): 74–78. doi: [10.1109/LGRS.2013.2246538](https://doi.org/10.1109/LGRS.2013.2246538).
 - [5] ZHAO An, FU Kun, SUN Hian, *et al.* An effective method based on acf for aircraft detection in remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(5): 744–748. doi: [10.1109/LGRS.2017.2677954](https://doi.org/10.1109/LGRS.2017.2677954).
 - [6] DIAO Wenhui, SUN Xian, ZHENG Xinwei, *et al.* Efficient saliency-based object detection in remote sensing images using deep belief networks[J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(2): 137–141. doi: [10.1109/LGRS.2015.2498644](https://doi.org/10.1109/LGRS.2015.2498644).
 - [7] REN S, HE K, GIRSHICK R, *et al.* In 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. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
 - [8] LIU Wei, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot multibox detector[C]. *Computer Vision and Pattern Recognition*, Boston, USA, 2015: 21–37. doi: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
 - [9] HUANG Gao, LIU Zhuang, and LAURENS van der Maaten. Densely connected convolutional networks[C]. *Computer Vision and Pattern Recognition*. Hawaii, USA, 2017: 567–576.
 - [10] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. *Computer Science*, 2014, 10(1): 123–132.
 - [11] ZHOU Bolei, KHOSLA A, LAPEDRIZA A, *et al.* Object detectors emerge in deep scene CNNs[J]. *Computer Science*, 2014, 16(2): 1205–1217.
 - [12] ERHAN D, SZEGEDY C, TOSHEV A, *et al.* In Scalable object detection using deep neural networks[C]. *Computer Vision and Pattern Recognition*, Columbus, USA, 2014: 2155–2162. doi: [10.1109/CVPR.2014.276](https://doi.org/10.1109/CVPR.2014.276).
 - [13] GIRSHICK R. Fast R-CNN[J]. *Computer Science*, 2015, 4(1): 1440–1448.
 - [14] EVERINGHAM M, GOOL L V, WILLIAMS C K I, *et al.* The pascal Visual Object Classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303–338. doi: [10.1007/s11263-009-0275-4](https://doi.org/10.1007/s11263-009-0275-4).
- 郭 智: 男, 1975年生, 研究员, 研究方向为地理空间信息综合处理与应用。
宋 萍: 女, 1991年生, 硕士生, 研究方向为机器学习与遥感图像智能解译。
张 义: 男, 1987年生, 助理研究员, 研究方向为阵列信号处理。
闫梦龙: 男, 1985年生, 副研究员, 研究方向为机器学习与遥感图像智能解译。
孙 显: 男, 1981年生, 副研究员, 研究方向为机器学习与遥感图像智能解译。
孙 皓: 男, 1984年生, 副研究员, 研究方向为机器学习与遥感图像智能解译。