

基于 SSD 框架的红外弱小目标检测技术研究

马玉莹,黄成章,黄静颖,王伟丞
(华北光电技术研究所,北京 100015)

摘要:针对红外序列图像中点状机动目标的检测问题,利用卷积特征对图像具有更细致的描述能力、可以有效地感知弱小目标与干扰的差异的特点,提出了一种基于 SSD 框架的红外弱小目标检测方法。本文设计了针对弱小目标的 SSD 检测框架,且克服了传统方法缺少大量实验数据的困难,利用真实红外图像数据库进行了仿真验证,实验结果表明,该方法对多种不同场景中的信噪比较低的目标具有良好的检测效果和虚警抑制能力。

关键词:弱小目标检测;卷积神经网络;杂波抑制

中图分类号:TN391.41 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2021.10.013

Research of infrared dim and small targets detection technology based on SSD framework

MA Yu-ying, HUANG Cheng-zhang, HUANG Jing-ying, WANG Wei-cheng
(North China Research Institute of Electro-Optics, Beijing 100015, China)

Abstract: Aiming at the detection problem of point-like maneuvering targets in infrared sequential images, an infrared dim and small targets detection method based on SSD framework is proposed, which can effectively perceive the difference between small and small targets and interference by taking advantage of the convolutional feature, which has a more detailed description ability to the images. In this paper, the SSD detection framework for dim small targets is designed, and it overcomes the difficulty that traditional methods lack a large amount of experimental data. Simulation verification is carried out by using real infrared image database. Experimental results show that this method has good detection effect and false alarm suppression ability for targets with low SNR in various scenes.

Keywords: dim and small target detection; convolutional neural network; clutter suppression

1 引言

深度学习是机器学习算法中较为复杂的一类方法,它的引入使学术研究与工程应用更加接近于人们盼望达到的人工智能的目标。深度学习的概念由 Geoffrey Hinton 等人于 2006 年率先提出,通过建立真正深度神经网络的训练机制来模拟人脑的学习过程,并希望以模拟人脑的多层抽象机制来实现对图

像、语音及文本等数据的抽象表达,从而自动完成特征提取过程,减少人为干预。深度学习中涉及的理论众多,网络及其丰富,有众多的应用领域和各自不同的适应方法,而卷积神经网络作为其中的一个重要分支,在基于深度学习的目标检测技术中,主要应用的便是深度学习中卷积神经网络的相关部分。卷积神经网络(Convolutional Neural Network, CNN)是

深度学习中非常重要的一个网络模型^[1],被用来设计解决图像识别等问题,能够避免复杂的特征提取工作。

随着光电探测系统对探测距离的要求日渐提高,从获取红外目标图像的过程来看,在目标与探测器距离较远时的一段时间内,开始时目标在成像平面上所占的像素很少,表现为一个对比度非常低的亮点,随着距离的减小,通常在图像中呈现为点状的动目标。在该种情形下的目标搜索检测任务的难度体现在很难用传统的特征描述方法去提取弱小目标的特征。随着卷积神经网络技术在其他领域取得的巨大突破,越来越多的人开始研究基于卷积神经网络的弱小目标探测技术^[2],但无论是传统图像处理算法设计,还是基于卷积神经网络的网络模型训练,都离不开图像数据,图像数据库成为制约卷积神经网络算法性能的关键因素之一。由于一切物体都会自发的辐射红外信号,光电系统中能接收到视场内一切物体产生的背景辐射信号,背景辐射信号中包含大量的“杂波”信号,通常认为“杂波”是场景中的干扰目标探测与跟踪性能的物体产生的“信号”,“杂波”已经成为制约光电系统性能的一个瓶颈,如何对“杂波”引起的虚警进行抑制,对提升光电系统的检测指标起着至关重要的作用。

鉴于以上原因以及在红外图像中,目标与周围的背景相比没有明显的亮度,且不能依赖于目标的灰度信息、目标的形状和纹理等信息来区分目标与杂波,考虑到这些因素给检测带来的困难,本文提出了一种基于深度学习方法(SSD 框架)的红外弱小目标检测技术来解决红外弱小目标的远距离探测和虚警抑制问题。

2 针对弱小目标的 SSD 检测框架

深度学习的核心思想体现在“非监督特征学习”(Unsupervised Feature Learning, UFL)。深度学习不需要通过人工设计特征提取算法的传统方式,而是让机器在无监督的条件下提取和选择特征,将基于特征表达的图像识别等任务统一在同一个理论框架下^[3]。基于深度学习的目标检测框架 SSD (single shot multibox detector)^[4]是目前的主流检测框架之一,该框架使用了特征金字塔结构进行目标检测,检测时在不同卷积层输出的特征上同时进行

Softmax 分类和位置回归,该算法检测时综合利用了浅层和深层的特征信息,因此对弱小目标能够有较好的检测精度。但是,红外场景中的弱小目标与其他场景弱小目标相比,具有目标更小,更难以分辨的特性,然而现有的 SSD 网络主要是基于 VGG_16 网络,网络层数较深,在卷积的过程中,红外小目标的特征信息容易丢失,红外小目标检测效果仍然不够理想。

针对上述问题,本文提出了一种基于双阈值多尺度的 SSD 网络框架。该网络框架主要针对红外图像中的弱小目标检测需求进行了两方面的改进:第一,使用网络层数较少的 AlexNet 替换 SSD 中的 VGG_16,并且在卷积的过程中使用 3×3 的卷积核替换 AlexNet 中较大的卷积核,使得网络具有一定的复杂度,整体的感受野又不会太大,在卷积的过程中不至于丢失弱小目标的特征信息;第二,借鉴了从粗粒度到细粒度的二级检测思想,设计了基于双阈值的粗粒度检测器和细粒度检测器。在目标检测任务中,对单阈值的设定往往要求比较高,阈值设置过高会漏掉许多真目标,导致检测率降低,阈值设置过低则会混入假目标,导致虚警率高。本文设置了两个阈值 p_1, p_2 (其中 $1 > p_1 > p_2 > 0$),用于指导粗粒度检测器和细粒度检测器,缓解上述问题。

2.1 粗粒度检测器

该粗粒度检测器主要是基于 SSD 进行改进的,原始的 SSD 网络结构如图 1 所示,VGG_16 网络结构作为基础,对 VGG_16 conv4_3 层及后续额外的网络层的输出特征图进行多尺度的目标检测。本文方法主要针对红外弱小目标,目标比一般小目标更小,VGG_16 网络对于该类型目标来说,网络太深,随着卷积层数的加深,目标的特征信息可能丢失,因此本文方法使用卷积层数较少的 AlexNet 替换 VGG_16,并且在卷积的过程中使用 3×3 的卷积核替换 AlexNet 中较大的卷积核,同时减少后续额外的网络层数,具体网络结构如图 2 所示。在检测过程中,将待检测原始红外图像送入粗检测器中,进行检测,当检测结果的置信值高于设定的阈值 p_1 ,直接标记该检测结果;置信值低于阈值 p_2 时,则判定该区域无目标;置信值大于 p_2 且小于 p_1 时,则送入细粒度检测

器做进一步更精细化的检测。

2.2 细粒度检测器

该细粒度检测器的网络结构与粗粒度检测器相同。对于粗粒度检测器中,检测结果置信度均小于 p_1 并大于 p_2 的区域可能存在着目标,该目标可能由于目标太小或者受到部分遮挡等因素,识别的确定性降低。本文将该区域裁剪出来,使用双三次上采

样的超分辨率技术^[5]提升该区域的分辨率,然后送入细粒度检测器中,做进一步的目标检测,当检测结果的置信值高于设定的阈值 p_1 ,则标记该检测结果,最后的检测结果综合了粗粒度检测器和细粒度检测器的结果。该操作仅对粗粒度检测器中,不易检测的区域进行更进一步的检测,仅增加了少量的计算量。

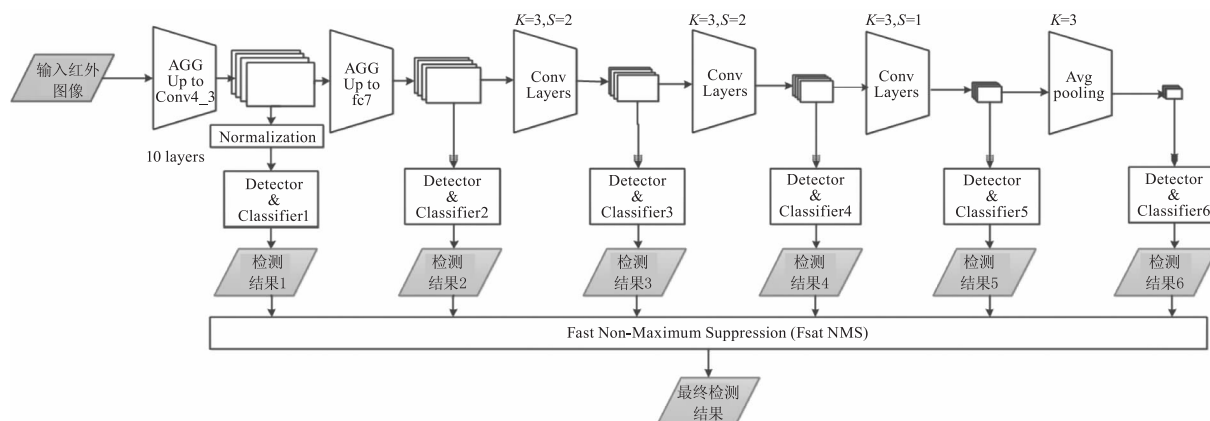


图 1 SSD 网络结构图

Fig. 1 SSD network structure diagram

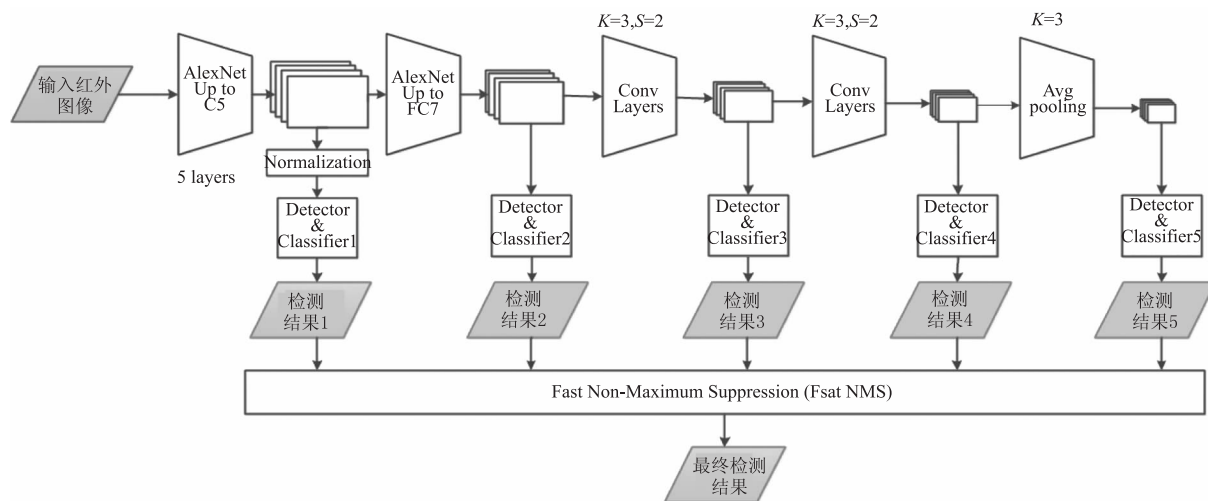


图 2 本文方法网络结构图

Fig. 2 Network structure diagram of the method presented in this paper

3 红外弱小目标检测系统

红外弱小目标检测算法的整体流程如图 3 所示:粗粒度检测器对原始红外图像进行检测,当检测置信度大于阈值 p_1 时,直接判定该类目标;当检测结果的置信度小于阈值 p_1 且大于阈值 p_2 时,则对该区域进行上采样,提高该区域的分辨率,然后送入细粒度检测器中,进行更进一步的检测。

3.1 杂波抑制

为了提高目标检测的准确率,该方法在目标检测前对图像进行了噪声抑制,减少噪声干扰导致的虚警(如图 3 所示)。红外图像中的盲闪元在空域中表现为孤立的点^[6],而目标距离较远时,在图像中只有一到几个像素。因此,在目标的检测率较高时,会出现大量因盲闪元引起的虚警。传

统的空域滤波算法,如中值滤波,虽然能去除孤立的噪声点,较好地保持物体边缘,但是尺寸特别小的目标也会被滤除,从而降低小目标的检测率。为了减小盲闪元对目标检测性能的影响,首先采用了自适应平滑滤波,通过自适应平滑滤波,使得图像的灰度值更加均匀,避免中值滤波的过程中,

将小目标当成噪声去除,该自适应平滑滤波的操作方式如下:

$$f(x, y) = \sum_{m, n=-P}^P K_{x, y}(m, n) f(x + m, y + n)$$

其中, $K_{x, y}$ 表示以 (x, y) 为中心的自适应平滑滤波核;滤波核大小为 $(2P + 1) \cdot (2P + 1)$ 。

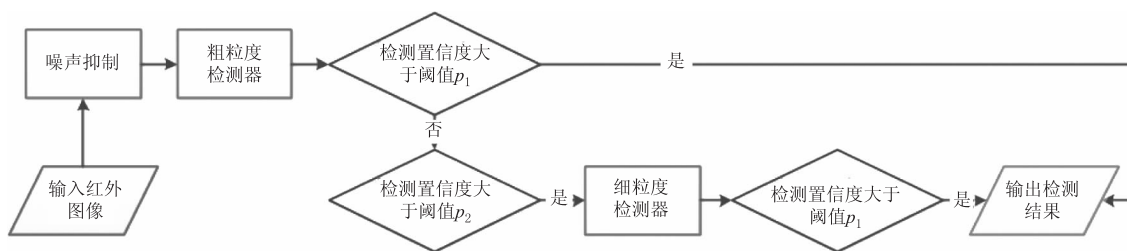


图3 整体的检测算法框架流程

Fig. 3 General framework flow of detection algorithm

3.2 训练损失函数

本文方法的损失函数与 SSD 算法一致,分成两个部

分,一部分计算通过 Softmax 类别的置信度损失,另一部分通过 SmoothL1 计算位置回归。具体算法流程如下所示。

算法:基于双阈值多尺度的 SSD 网络框架训练流程

输入:待处理红外图像,阈值 p_1 和 p_2 ,批处理量 n_b

输出:含目标信息的图像

1. 对红外图像数据集进行目标检测前噪声抑制预处理得到数据集 $\hat{X}$;
2. For num = 1; num < 迭代次数 do
3. 从训练集中随机采样待检测图像 $\hat{X} = \{\hat{x}^{(1)}, \dots, \hat{x}^{(n_b)}\}$;
4. 获取待检测图像对应的类别和位置标签 $C = \{c^{(1)}, \dots, c^{(n_b)}\}, L = \{l^{(1)}, \dots, l^{(n_b)}\}$;
5. 将待检测图像送入粗粒度检测器中,将检测结果按照阈值大于 p_1 和阈值大于 p_2 且小于 p_1 区域分别分成两组 g_1 和 g_2 ;
6. 对 g_2 中包含的区域进行双三次上采样,送入到细粒度检测器中进行检测;
7. 计算训练的损失函数,并通过梯度下降方式更新粗粒度和细粒度检测器的参数。

4 实验结果及分析

4.1 评价指标分析

在多目标分类器的判别中,设目标的种类数为 n 。对单种目标的判别仍然遵循每一种假设有两种结果的四中可能性,即设 $D_i^j (j = 1, 2, \dots, n)$ 表示一种目标 j 选择假设 H_i^j 为真,任何二元假设试验问题中,做判别时要考虑四种可能性^[7]:

A) H_0^j 假设为真,判别为 D_0^j ;

B) H_0^j 假设为真,判别为 D_1^j ;

C) H_1^j 假设为真,判别为 D_0^j ;

D) H_1^j 假设为真,判别为 D_1^j ;

A 和 D 对目标 j 选择正确;B 称为第一类错误,叫做虚警(真实情况没有目标却判别为有目

标);C 称为第二类错误,叫做漏报(真实情况有目标而判别为没有目标)。除此之外,在多目标识别中还有将目标 D_i^j 识别为 $D_i^k (k = 1, 2, \dots, n, k \neq j)$ 的错误判断。

本文选取检测率和虚警率作为考察算法性能的指标,设目标 Z^j 在判别域 Z_0^j 和 Z_1^j 上的概率密度函数分别为 $f(z^j | H_0^j)$ 和 $f(z^j | H_1^j)$,理论上则有:

虚警率:

$$P_f = \sum_{j=1}^n P(D_1^j | H_0^j) = \sum_{j=1}^n \int_{Z_1^j} f(z^j | H_0^j) dz$$

检测率:

$$P_d = \sum_{j=1}^n P(D_1^j | H_1^j) = \sum_{j=1}^n \int_{Z_1^j} f(z^j | H_1^j) dz$$

红外弱小目标检测算法的性能指标主要包括检测概率和虚警率。对于给定的算法,检测概率和虚警率与检验图像之间的关系非常相关。传统的理论推导建模方法完全采取数学推导对目标特性与算法的检测概率与虚警率的关系进行理论建模。由于在建模之前作了大量的假设和简化操作,这种方法在准确性上与实际情况相差的较远。实际上,检测算法的性能不仅仅依赖于目标特性,而且与红外图像背景特性也有一定的关联。

4.2 仿真结果分析

为了验证算法的有效性,研究过程中构建了仿真数据库,对确定的目标和干扰进行了半自动人工标注,获取了大量的目标及干扰样本数据集。为了提高模型的泛化能力,样本的选取综合考虑了各种环境因素(纯空背景、云层背景及地面背景)影响,从各种场景中等比例的抽取了正样本及负样本,采用留出法构建了训练集(70%)和测试集(30%)。为了真实有效的评估算法性能,本文利用具有代表性的环境因素(天空背景、云层背景、地面背景等)的图像进行测试,将传统算法的检测概率和本文算法的检测概率进行了测试统计和对比。在这里定义检测概率:

$$P_d = \frac{N_a}{N_b}$$

其中, N_a 为图像序列中检测到真实目标个数; N_b 为图像序列中实际目标个数。

目标分类中,研究者们关心的是存在的目标的识别效果,识别率一般指检测率,在上面的 3 种代表性场景中目标检测率如表 1 所示,在确保虚警率满足系统指标要求的前提下,算法的检测率实现了最大限度的提升,其中天空背景的检测率高达 99%。

表 1 不同复杂度场景的信噪比和检测概率
Tab. 1 SNR and detection probability of scenarios with different complexity

序号	背景类型	图像信噪比	传统算法检测概率	本文算法检测概率
场景 1	天空	2.58	0.73	0.99
场景 2	地面	2.76	0.66	0.94
场景 3	云层	2.43	0.61	0.85

虚警的抑制效果以天空背景(场景 1)、地面背

景(场景 2)、云层背景(场景 3)这三类代表性场景为例,利用传统目标检测算法仿真效果如图 4 所示,算法中加入深度学习模型后达到的效果如图 5 所示。其中,天空背景的复杂度是最低的,检测的难度也相对效低,地面和云层背景复杂度较高,传统方法中预处理等杂波抑制过程的效果有限,还是会有存在较多的虚警。

从上述图像的处理结果来看,算法中未加深度学习模型时,在云层边缘等梯度较大的区域,最容易产生虚警;算法中加入深度学习模型后,模型会根据每个像素点的局部区域计算置信度,只有置信度大于设定阈值时,才会将其作为候选目标,进行后续处理,可以较好地抑制虚警。

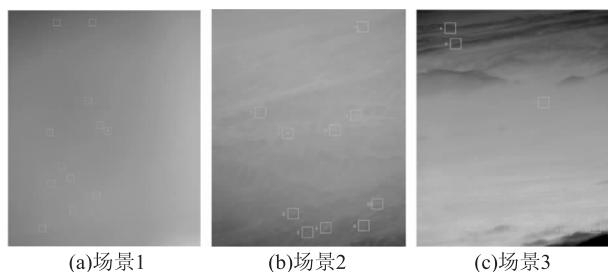


图 4 传统目标检测算法仿真效果

Fig. 4 Simulation effect of traditional target detection algorithm

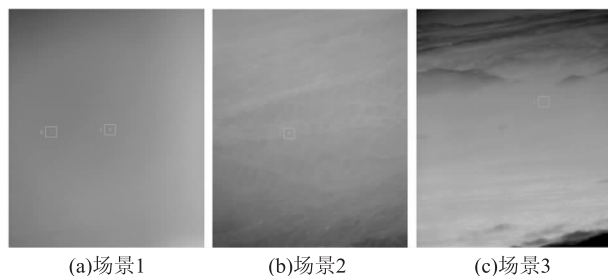


图 5 本文目标检测算法仿真效果

Fig. 5 Simulation effect of target detection algorithm in this paper

5 结 语

本文的创新点体现在以下两个方面:利用真实红外图像建立了自己的数据库,将深度学习技术应用到了红外弱小目标检测问题中;本文设置了两个阈值用于指导粗粒度检测器和细粒度检测器,缓解了目标检测任务中,对单阈值的设定要求高的问题。基于真实红外序列图像的实验效果验证了本文的提出的方法有效性,在很大程度上提高了红外弱小目标的检测概率的同时控制了虚警率,在光电系统目标探测过程中具有很强的实用性。

参考文献:

- [1] Chang Liang, Deng Xiaoming, Zhou Mingquan, et al. Convolutional neural networks in image comprehension [J]. Acta Automatica Sinica, 2016, 42(9): 1300 – 1312. (in Chinese)
常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络[J]. 自动化学报, 2016, 42(9): 1300 – 1312.
- [2] Zhou Feiyan, Jin Linpeng, Dong Jun. A review of convolutional neural networks [J]. Chinese Journal of Computers, 2017, 40(6): 1229 – 1251. (in Chinese)
周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229 – 1251.
- [3] Zeiler M D, Fergus R. Visualizing and understanding convolutional neural networks [C]//European Conference on Computer Vision, 2013.
- [4] Liu Wei, et al. SSD: single shot multibox detector [C]//Proc of European Conference on Computer Vision. Cham: Springer, 2016: 21 – 37.
- [5] Lin Tsuang Yi, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [6] Gao Mingfei, et al. Dynamic zoom-in network for fast object detection in large images [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018.
- [7] Ma Chunting, Zheng Jian, Chen Donggen, et al. Evaluation index of multi-target recognition for ground battlefield reconnaissance system [J]. Journal of Detection & Control, 2006, 28(1): 6 – 9. (in Chinese)
马春庭, 郑坚, 陈东根, 等. 地面战场侦察系统多目标识别的评价指标[J]. 探测与控制学报, 2006, 28(1): 6 – 9.