

# 空基下视多角度红外目标识别 \*

刘 彤<sup>1,2,3</sup>, 杨德振<sup>1,4</sup>, 宋嘉乐<sup>1,2,3</sup>, 傅瑞罡<sup>3</sup>, 何佳凯<sup>1</sup>

(1. 华北光电技术研究所 机载探测中心, 北京 100015; 2. 中国电子科技集团公司电子科学研究院, 北京 100015;  
3. 国防科技大学 电子科学学院 ATR 重点实验室, 湖南 长沙 410073;  
4. 北京真空电子技术研究所 微波电真空器件国家重点实验室, 北京 100015)

**摘要:** 为实现空基下视红外目标的快速高精度识别, 提出了一种单阶段的空基下视多角度红外目标识别算法。首先使用 Darknet-53 结合 SPP 模块对红外目标进行特征提取, 使局部特征与全局特征融合, 提高特征图表达能力, 最后借鉴 RetinaNet 中的 Focal loss 锁定目标的检测框, 同时得出目标类型及检测精度。针对现有数据集多为平视, 且视角单一的缺陷, 使用复合翼无人机分别从不同高度和角度采集红外图像, 构建多尺度下视红外目标数据集, 在 PyTorch 架构上实现并进行性能验证实验, 所提算法对分辨率 640×512 的下视红外图像中目标识别的 mAP 达到 91.74%, 识别速度为 33 f/s, 满足空基平台前端的在线识别需求, 且在公开红外船舶数据集上也具有较好的识别结果。实验表明该算法在保证精度的基础上满足实时性的要求, 为后续用于复合翼无人机上的多尺度目标实时识别提供了理论技术。

**关键词:** 下视红外目标; 单阶段; 复合翼无人机; SPP; Focal loss; 多尺度目标

中图分类号: TP391.9

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.212449

中文引用格式: 刘彤, 杨德振, 宋嘉乐, 等. 空基下视多角度红外目标识别[J]. 电子技术应用, 2022, 48(7): 131–139.

英文引用格式: Liu Tong, Yang Dezhen, Song Jiale, et al. Air-based downward-looking multi-angle infrared target recognition[J]. Application of Electronic Technique, 2022, 48(7): 131–139.

## Air-based downward-looking multi-angle infrared target recognition

Liu Tong<sup>1,2,3</sup>, Yang Dezhen<sup>1,4</sup>, Song Jiale<sup>1,2,3</sup>, Fu Ruigang<sup>3</sup>, He Jiakai<sup>1</sup>

(1. Airborne Detection Center, North China Research Institute of Electro-optics, Beijing 100015, China;

2. China Academic of Electronics and Information Technology, Beijing 100015, China;

3. ATR Key Laboratory, National University of Defense Technology, Changsha 410073, China;

4. National Key Laboratory of Science and Technology on Vacuum Electronics, Beijing Vacuum Electronics Research Institute, Beijing 100015, China)

**Abstract:** In order to realize the rapid and high-precision recognition of air-based downward-looking infrared targets, a single-stage space-based down-looking multi-angle infrared target recognition algorithm is proposed. Firstly, use Darknet-53 combined with SPP module to perform feature extraction on infrared targets, to fuse local features and global features to improve the expression ability of feature maps, and finally use Focal loss in RetinaNet to lock the detection box of the target, and at the same time obtain the target type and detection accuracy. Aiming at the defect that the existing data sets are mostly head-ups and single viewing angles, a composite-wing drone was used to collect infrared images from different heights and angles, and a multi-scale down-view infrared target data set was constructed, which was implemented and verified on the PyTorch architecture. The proposed algorithm achieves 91.74% of the mAP of the downward-looking infrared target, the recognition speed is 33 FPS, and it also has a good recognition result on the public infrared ship data set. The experiments show that the algorithm meets the real-time requirements on the basis of ensuring accuracy, and provides theoretical technology for subsequent real-time recognition of multi-scale targets on compound-wing UAVs.

**Key words:** downward-looking infrared target; one stage; compound wing drone; SPP; Focal loss; multi-scale target

## 0 引言

采用多个复合翼无人机集群的空基红外目标探测识别系统进行协同态势感知是地面远程红外制导的有

效辅助手段。其中舰船和车辆目标的智能识别算法作为当代陆海防务系统中关键技术, 对于保障我国的国防安全具有重要的应用价值。红外图像具有对比度高、作用距离远、穿透力强等特点, 热成像不需要借助外界光源, 隐蔽性好, 可以全天时工作。目前多数车载系统的视角

\* 基金项目: 国家自然科学基金青年科学基金(62001482)

为平视,平视视场受限,而使用机载系统可以获得更大的下视视场角,在实战中能有效探测和拦截低空突防的威胁目标。

传统目标识别算法难以对下视视场角的图像进行特征提取,作为端到端学习的深度学习,能够自动提取图像特征,具有较好的目标识别能力。北京理工大学的王旭辰等<sup>[1]</sup>提出基于深度学习的无人机载多平台目标检测算法,使用 Darknet-53 网络作为检测器,在公开数据集 UAV123 和实测数据集中进行检测,验证得到该算法在视角旋转、目标尺度变化以及障碍物遮挡下能进行稳定检测。但文中仅使用了 YOLOv3 目标检测算法,并未对其进行改进,检测精度没有得到提高。电子科技大学的刘瑞<sup>[2]</sup>针对空中目标存在目标尺度及疏密程度变化大、存在重叠、遮挡等问题,提出四级复杂度的航空图像目标检测算法,采用复合扩张主干网深度和宽度的方法构建出四级复杂度的主干网络,再将主干网分别与 FPN+PAN 网络、输出头网络相结合,得到空中目标检测算法。在 VisDrone-DET2020 训练集下对提出的网络进行训练,将算法的 mAP@[.5:.95]累计提升了 0.65%,mAP@0.75 累计提升了 1.41%。但仅在公开数据集上进行训练与检测,该公开数据集仅为可见光数据集,不具备红外数据集所具备的优点。严开忠等<sup>[3]</sup>针对小型无人机载平台算力受限、检测速度慢的问题,提出了一种改进 YOLOv3 的目标检测算法,引入深度可分离卷积对主干网络进行改进,降低网络的参数和计算量,从而提高检测速度。此外使用 K-means 生成先验框的初始聚类中心,在边界框回归中使用 CIoU 损失,同时将 DIoU 与 NMS 相结合,提高算法的检测精度。在自定义数据集中的 mAP 为 82%,检测速度从 3.4 f/s 提高到 16 f/s。但算法的检测精度和速度仍有待提高。上海交通大学的朱壬泰等<sup>[4]</sup>针对目前深度学习中多目标检测算法占用资源量大,无法在中小型无人机平台上实时运行的问题,提出了基于卷积神经网络计算优化的算法。采用深度可分离卷积对计算量进行优化,将主干网 Resnet18 中的卷积层替换为深度可分离卷积,对改进的算法在公开数据集 PASCAL VOC 2007 进行验证,得到在检测精度不变的条件下,检测速度达到 56 f/s。但该算法在航拍数据集上的检测精度由于与公开数据集分布的差异有所下降,对航拍目标的适应性不强。周子衿<sup>[5]</sup>针对深度学习网络结构复杂,机载平台计算资源有限,以及航拍视角中小目标数量大,难以对图像特征进行提取,容易出现漏检和错检现象,对 YOLOv3 网络进行轻量化处理,在网络稀疏化训练后进行 BN 层的通道剪枝,此外使用 K-means++ 算法对先验框进行重定义,将改进后的算法在自定义的 DOTA-like 数据集上进行测试,算法权重模型大小下降 98.7%,使得推理时间加快了 60.5%,检测速度提高了 32.9%,检测精度提高 1.14%。但无人机的飞行高度较低,局限于超低空域附近,所看到的视场角受限。

上述文章中均不存在多角度的空基下视红外目标,且该类目标所在环境复杂多样,为保证较高的精度和实时性,提出了一种基于单阶段的多角度空基下视红外目标识别算法。主要研究内容包括单阶段目标检测算法、针对空基下视多角度红外目标所提出的改进、多角度下视红外数据集的构建以及所提改进算法与单阶段算法的对比实验。

## 1 单阶段算法

根据深度学习算法是否需要经过候选框的步骤,将目标检测算法分为单阶段和双阶段两类。双阶段检测算法虽然检测精度较高,但运行速度慢,无法满足多角度空基下视红外目标实时识别的需求<sup>[6]</sup>。常用的单阶段算法有 SSD<sup>[7]</sup>、RetinaNet、YOLO 等,SSD 算法检测精度较低,RetinaNet 算法检测精度得到了提升,但检测速度无法达到实时性需求。YOLO 算法在拥有较高检测精度的同时兼顾检测速度。YOLOv3 借助残差网络思想<sup>[8]</sup>,将 YOLOv2 特征提取网络的 Darknet-19 改为 Darknet-53<sup>[9]</sup>。主要由  $1 \times 1$  的卷积层和  $3 \times 3$  的卷积层组成该特征提取网络。为防止过拟合,在每个卷积层之后加入一个批量归一化层和一个 Leaky ReLU。Darknet-53 中的基本单元 DBL 包含卷积层、批量归一化层和 Leaky ReLU 三个组成部分。Darknet-53 中共有 53 个 DBL 基本单元。res unit 为残差单元,输入通过两个 DBL,再与原输入相加。该单元能够让网络提取到更深层次的特征,并且能避免出现梯度的消失或梯度爆炸。

为实现多尺度特征融合的目的,把 Darknet-53 的中间层和后面某一层的上采样进行拼接,对张量的维度进行扩充。从而得到三种不同尺度的输出 Y1、Y2、Y3。Darknet-53 相比于 Darknet-19,不再采用最大池化层,而是使用步长为 2 的卷积层进行下采样。将 BN 层和一个 Leaky ReLU 加在每个卷积层后,以防止过拟合现象的发生。同时,为提取到更深层次的特征,引入残差网络的思想,可以同时避免梯度消失或梯度爆炸。为了能够进行多尺度特征融合,将网络中间层和后面某一层的采样结果进行张量拼接。

在多尺度预测方面,YOLOv3 选择三种不同形状的 Anchors,每种 Anchors 拥有三种不同的尺度,得到共计 9 种不同大小的 Anchors,达到预测多尺度目标的目的。

在损失函数方面,YOLOv3 相比于 YOLOv1 的损失函数,其位置损失部分仍使用 sum-square error 的损失计算方法。但对于置信度损失和类别预测将原来的 sum-square error 计算方法改为交叉熵损失计算方法。

## 2 YOLOv3-SF

### 2.1 网络结构的改进

由于复合翼无人机在不同时刻采集红外图像时所处的高度不同,导致获得的图像尺寸有所差异。图像空间金字塔池化(Spatial Pyramid Pooling, SPP)<sup>[10]</sup>结构能够

解决因图像区域裁剪、缩放一系列操作导致的图像失真问题，并能解决卷积神经网络对图像特征的重复提取，极大地提升了产生候选框的速度，同时节约计算成本。SPP 模块结构如图 1 所示。

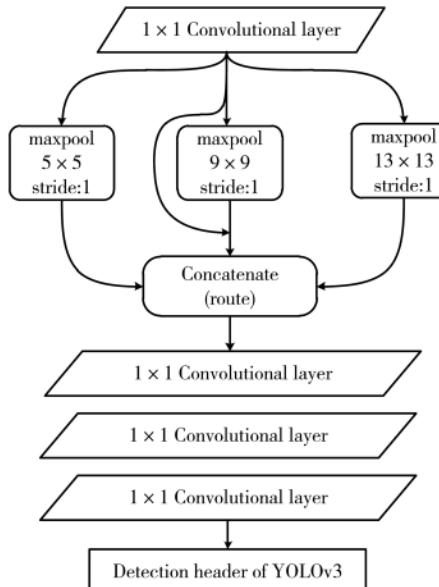


图 1 SPP 模块结构图

SPP 模型由 kernel size 为  $5 \times 5$ 、 $9 \times 9$ 、 $13 \times 13$  的最大池化和一个跳跃连接四个并行分支组成。借鉴空间金字塔池化思想，对 YOLOv3 网络结构进行改进。在原 YOLOv3 网络的第 5、6 层卷积之间加入一个 SPP 模块，不同的池化操作组成该模块。改进后的网络结构如图 2 所示。

通过检测头前的第 5 卷积层和第 6 卷积层之间集成的 SPP 模块获取网络中的 SPP，经过 SPP 模块池化后的特征图重新按通道进行拼接传入到下一层侦测网络中。该模块借鉴了空间金字塔思想，利用 SPP 模块实现局部特征和全局特征，特征图经局部特征和全局特征融合后，能够丰富特征图的表达能力，有利于提升待检测图像中目标大小差异较大情况下的检测精度。使用 SPP 模

块优化主干网络 Darknet-53 的算法用 YOLOv3-S 表示。

## 2.2 损失函数的改进

在所得的空基下视多角度红外图像中，目标较背景的占比很小，存在正负样本不平衡的现象。在训练密集目标检测器时出现严重的前景-背景类别不平衡，一步检测算法精度无法超过两步网络。但 RetinaNet 的提出使得一步网络的检测精度能够首次超越两步网络，因此借鉴 RetinaNet 的思想对 YOLOv3 进行改进，使得改进后的算法在拥有较快检测速度的同时检测精度也得到了相应的提高。原有交叉熵损失无法抵抗类别极不平衡，易导致分类器训练失败，因此使用 Focal loss<sup>[11]</sup>代替 YOLOv3 中使用的交叉熵误差。在正负样本匹配方面，针对每一个 Anchor 于事先标注好的 GT 进行匹配，计算其 IOU。若某一个 Anchor 与其 GT box 的  $\text{IOU} \geq 0.5$ ，则将 Anchor 标记为正样本；若某一个 Anchor 与所有的 GT box 的 IOU 值都  $< 0.4$ ，则将其标记负样本；若  $\text{IOU} \in [0.4, 0.5)$ ，则将其舍弃。对于二分类而言普通的交叉熵损失(CE)计算公式如式(1)所示。

$$\text{CE}(p, y) = \begin{cases} -\log(p), & y=1 \\ -\log(1-p), & \text{其他} \end{cases} \quad (1)$$

$p_t$  定义如式(2)所示。

$$p_t = \begin{cases} p, & y=1 \\ 1-p, & \text{其他} \end{cases} \quad (2)$$

将 CE 进一步简化得到式(3)。

$$\text{CE}(p, y) = \text{CE}(p_t) = -\log(p_t) \quad (3)$$

引入针对于正样本而言的超参数  $\alpha \in [0, 1]$  和针对于负样本而言的参数  $1-\alpha$ ，以此平衡正负样本的权重，CE 的表达式可改写为式(4)。

$$\text{CE}(p_t) = -\alpha_t \log(p_t) \quad (4)$$

$\alpha$  虽可以平衡正负样本的权重，但不能区分困难的样本和容易的样本。因此提出能够降低简单样本的权重，从而聚焦于一些难分的负样本。引入新元素  $(1-p_t)^\gamma$ ，定义 Focal loss 如式(5)所示：

$$\text{FL}(p_t) = -(1-p_t)^\gamma \log(p_t) \quad (5)$$

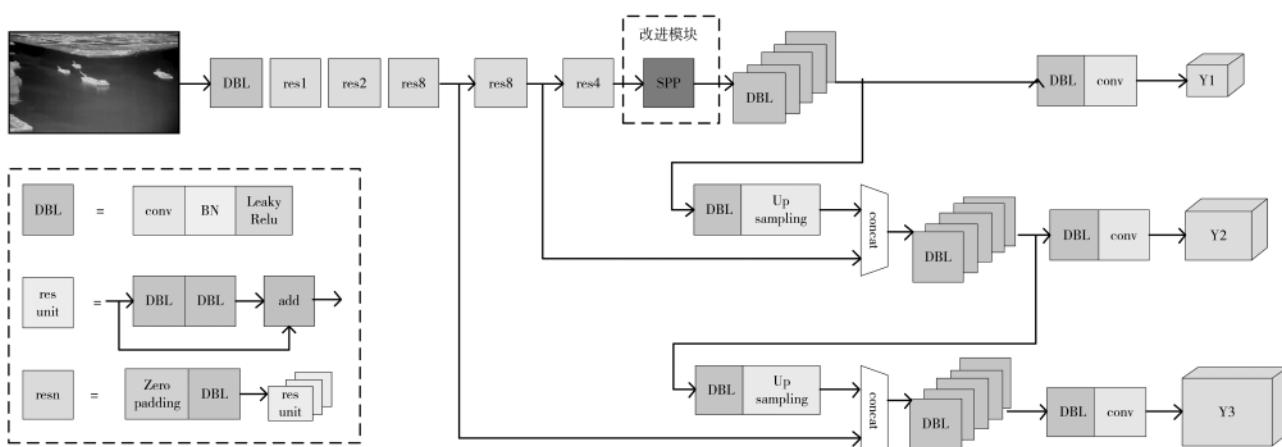


图 2 改进后的网络结构图

$(1-p_i)^\gamma$  能够降低易分样本的损失贡献。加入 $(1-p_i)^\gamma$  平衡的超参数 $\alpha_t$  得到 Focal loss 的最终形式如式(6)所示：

$$FL(p_i) = -\alpha_t(1-p_i)^\gamma \log(p_i) \quad (6)$$

将式(2)代入式(6)得式(7)：

$$FL(p) = \begin{cases} -\alpha(1-p_i)^\gamma \log(p_i), & \gamma=1 \\ -(1-\alpha)p^\gamma \log(1-p), & \text{其他} \end{cases} \quad (7)$$

Focal loss 与交叉熵对于难分样本和易分样本对比如表 1 所示。

表 1 Focal loss 与交叉熵在难分/易分样本对比

p	label	CE	FL	rate
0.9	1	0.105	0.000 26	400
0.968	1	0.033	0.000 008	3 906
0.1	0	0.105	0.000 79	133
0.032	0	0.033	0.000 025	1 302
0.1	1	2.3	0.466	4.9
0.9	0	2.3	1.4	1.6

其中 $p$  为预测目标概率；label 为对应的真实标签，1 表示正样本，0 表示负样本；CE 为针对二分类的交叉熵损失；FL 为 Focal loss；rate 为采用 CE 计算得到的损失和采用 Focal loss 计算所得到损失的比值。由表 1 可知第 1,2 行正样本预测概率为 0.9 和 0.968 属于易分样本， $p=0.9$  时，其 Focal loss 约为原来的 CE 损失的 1/400， $p=0.968$  时，其 Focal loss 约为原来的 CE 损失的 1/3906；第 3,4 行负样本预测概率为 0.1 和 0.032 也属于易分样本， $p=0.1$  时，其 Focal loss 约为原来的 CE 损失的 1/133， $p=0.032$  时，其 Focal loss 约为原来的 CE 损失 1/1302；第 5 行对应为正样本，网络预测概率为 0.1，该样本比较难学习，其 Focal loss 约为原来的 CE 损失的 1/4.9，第 6 行为负样本，网络预测概率为 0.9，同样为难分样本，其 Focal loss 约为原来的 CE 损失的 1/1.4。使用 Focal loss 后更加专注于训练难分的样本，而对于易分样本起到降低损失权重的作用。

总损失函数分为分类损失和回归损失两部分，见式(8)：

$$\text{Loss} = \frac{1}{N_{\text{POS}}} \sum_i L_{\text{cls}}^i + \frac{1}{N_{\text{POS}}} \sum_j L_{\text{reg}}^j \quad (8)$$

分类损失针对所有的正负样本计算，回归损失针对所有的正样本计算。其中 $L_{\text{cls}}$  为 Sigmoid Focal Loss， $L_{\text{reg}}$  为 $L_1$  损失， $N_{\text{POS}}$  为正样本个数， $i$  为所有的正负样本， $j$  为所有的正样本。使用 Focal loss 优化损失函数的 YOLOv3 算法用 YOLOv3-F 表示。同时使用 SPP 模块和 Focal loss 函数的算法用 YOLOv3-SF 表示。

### 3 实现与结果分析

#### 3.1 实验平台环境

系统由复合翼无人机作为光电探测设备的空中移动平台，通过图传设备进行无线数据传输，地面配备图像处理计算机<sup>[12]</sup>。训练网络模型使用硬件平台为 R7-4800H

锐龙处理器，GPU 采用 NVIDIA GeForce RTX 2060，采用 PyTorch 深度学习框架搭建该网络。

#### 3.2 红外图像采集及数据集构建

使用某型复合翼无人机搭载的多波段光电吊舱，采集大量海上舰船和陆地车辆红外目标，图像分辨率为 640×512，多尺度无人机图像采集系统现场如图 3 所示。



图 3 多尺度无人机图像采集系统现场图

复合翼无人机分别从低(200 m)、中(300 m)、高(500 m)不同高度对复杂背景中的车辆(其中车类包含小汽车、卡车、公交车、货车)、海上的游艇、货船等多种不同类型的目标俯拍，其中包含了大、中、小三种不同尺度的红外图像。按 8:1:1 的比例对数据集进行划分，为测试模型的鲁棒性，选取了一些复杂背景下不同运动速度的目标。

高度为 200 m 的红外大图像数据集中，车类、游艇和货船的下视角为 12°、16°、20°，训练集、验证集和测试集张数如表 2 所示。

表 2 高度为 200 m 红外图像数据集

类型	下视角/(°)	训练集/张	验证集/张	测试集/张
车	12	2 891	361	362
	16	2 576	322	322
	20	2 533	317	316
游艇	12	1 438	180	180
	16	1 390	174	174
	20	1 572	197	196
货船	12	1 245	156	155
	16	1 269	159	158
	20	1 046	131	131

高度为 300 m 的红外中型图像数据集中，车类、游艇和货船的下视角为 24°、28°、32°，训练集、验证集和测试集张数如表 3 所示。

高度为 500 m 的红外小图像数据集中，车类、游艇和货船的下视角为 35°、36°、40°，训练集、验证集和测试集张数如表 4 所示。

综上自定义红外数据集包括 29 853 幅红外舰船图像和 30 000 幅车辆红外图像。数据集的部分样本图像如图 4 所示。对所采集图像中的车辆和舰船目标使用 labelImg 标注软件进行标注，将标注好的带有车辆和舰船目标的红外图像送入 YOLOv3-SF 网络中进行联合训练。

表 3 高度为 300 m 红外图像数据集

类型	下视角/(°)	训练集/张	验证集/张	测试集/张
车	24	2 784	348	348
	28	2 982	373	373
	32	2 234	279	280
游艇	24	1 456	182	182
	28	1 293	162	161
	32	1 079	135	135
货船	24	1 547	193	194
	28	1 398	175	175
	32	1 187	148	149

表 4 高度为 500 m 红外图像数据集

类型	下视角/(°)	训练集/张	验证集/张	测试集/张
车	35	2 991	374	374
	36	2 687	336	336
	40	2 322	290	291
游艇	35	1 092	137	136
	36	1 563	195	196
	40	1 468	184	183
货船	35	1 351	169	169
	36	1 087	136	136
	40	1 399	175	175

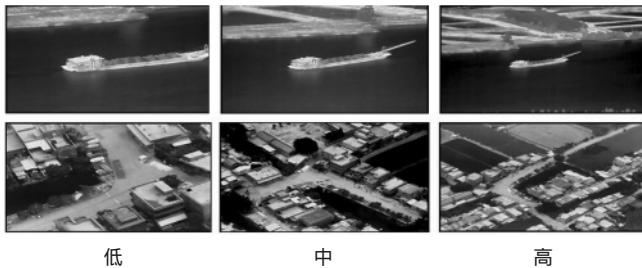


图 4 不同高度拍摄的红外图像

### 3.3 评价指标

在评价指标方面选取准确率(Precision)、召回率(Recall)、AP、mAP、FPS 作为该目标检测算法的评价指标<sup>[13]</sup>。TP、FN、TN、FP 的含义如表 5 所示。

表 5 正、负例含义表

符号	简称	含义	之和
TP	真正例	识别对了的正例	实际的正例数量
FN	伪负例	识别错了的负例	
TN	真负例	识别对了的负例	实际的负例数量
FP	伪正例	识别错了的正例	

注: TP 表示 True Positives, FN 表示 False Negatives, TN 表示 True Negatives, FP 表示 False Positives。正例为所关注的识别目标, 负例为正例以外的样本。

#### (1) 准确率(Precision)/查准率

识别对了的正例占识别出正例的比例, 其中识别出的正例等于识别对的正例加上识别错的正例, 公式如下:

$$P = \frac{TP}{(TP+FP)} \quad (9)$$

#### (2) 召回率(Recall)/查全率

识别对了的正例占实际总正例的比例, 其中实际总正例等于识别对的正例加上识别错的负例, 公式如下:

$$R = \frac{TP}{(TP+FN)} \quad (10)$$

#### (3) AP(Average Precision)

PR 曲线下的面积, 表示召回率 0~1 的平均精度。其中 PR 曲线下面积越大, 表示模型的性能越好。使用插值法计算 PR 曲线下的面积得到 AP, 公式如下:

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1\}} P_{\text{interp}}(r) \quad (11)$$

#### (4) mAP(mean Average Precision)

所有类别 AP 的平均值, 能衡量模型在所有类别上平均精度的好坏。该指标综合考虑了准确率 P 和召回率 R, 能够解决 P 和 R 的单点值局限性, 试算公式如下:

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (12)$$

### 3.4 实验结果及分析

设置初始学习率 learning\_rate=0.001, 迭代次数 steps=100, 训练过程中 batch 设置为 64, 输入该网络的图像尺寸为 416×416。在 YOLOv3-SF 的训练过程中训练损失(train loss)、验证损失(val loss)的变化曲线如图 5 所示, mAP 在不同模型下的变化曲线如图 6 所示。

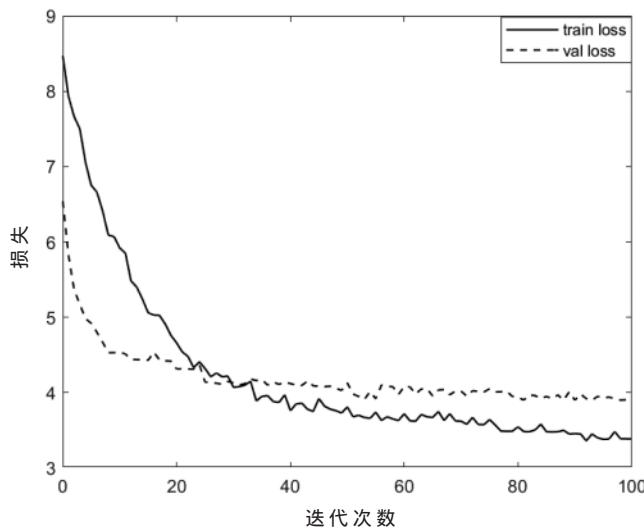


图 5 损失函数变化曲线

如图 5 所示损失函数总体呈下降趋势, 前 25 次迭代中训练损失快速下降, 最后趋于稳定。对于模型的识别精度, 以 mAP 作为评价指标, 如图 6 所示 YOLOv3-SF 算法的 mAP 最高, 且变化曲线更为稳定。

在自定义的红外图像测试集上分别选取低、中、高三种不同高度及下视角的舰船目标和车辆目标对 SSD

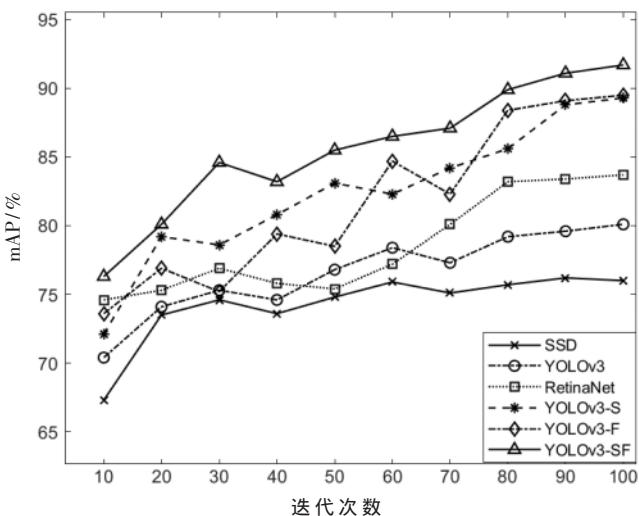


图 6 mAP 值在不同模型下的变化曲线

算法、YOLOv3 算法、RetinaNet 算法与 YOLOv3-S、YOLOv3-F、YOLOv3-SF 算法进行测试, 几种算法对比结果如图 7~图 10 所示。

由图 7~图 10 结果可以看到 SSD 和 YOLOv3<sup>[9]</sup>无法识别下视远距离小目标和被遮挡目标, 且 SSD 对于复杂

背景下的红外目标识别漏检率最高; RetinaNet<sup>[14]</sup>对下视远距离小目标的识别有了一定的提升, 但对遮挡目标的识别较差且存在一定的误检率; YOLOv3-SF 结合了 YOLOv3-S 和 YOLOv3-F 的优势, 识别效果得到较大提升, 对遮挡目标和小尺寸目标的识别效果最好。

表 6、表 7 分别给出了红外舰船目标和车辆目标在 6 种不同算法的输入图像尺寸和各类评价指标的检测结果。

在表 6、表 7 中的 AP 为 mAP 值; AP<sup>IoU=0.5</sup> 为 IoU=0.5 时的 mAP 值, 是 voc 的评判标准; AP<sup>IoU=0.75</sup> 为较为严格的 mAP 值, 能够反映算法框的位置精准程度; FPS 为每秒帧率, 即每秒内可以处理得到图片数量。

由各类算法的识别结果和表 6、7、8 可知, 在测试 3 组低、中、高红外舰船图像和红外车辆图像时, YOLOv3-SF 算法的识别精度最高, 识别速度较 RetinaNet 得到极大的提升, YOLOv3 算法识别精度较低, SSD 算法识别精度最低。说明 YOLOv3-SF 算法相比于原有 YOLOv3 算法在识别精度上得到较大提升, 可以在提高识别精度的同时兼顾识别速度, 从图 10 第三列可以看出在下视背景复杂度较高的场景应用具有一定优势。

此外将改进算法应用于公开红外船舶数据集 Maritime

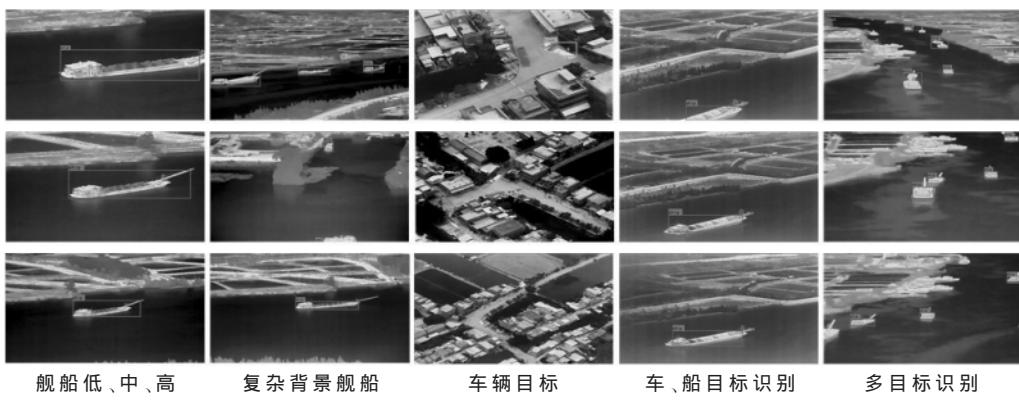


图 7 SSD 识别结果

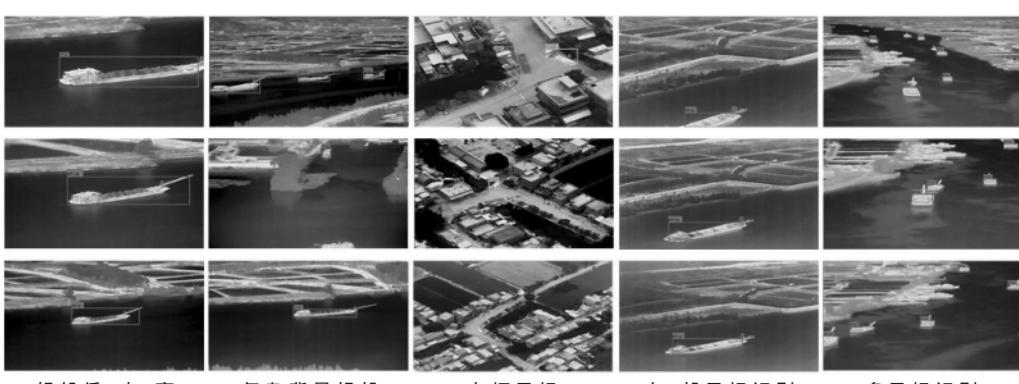


图 8 YOLOv3 识别结果

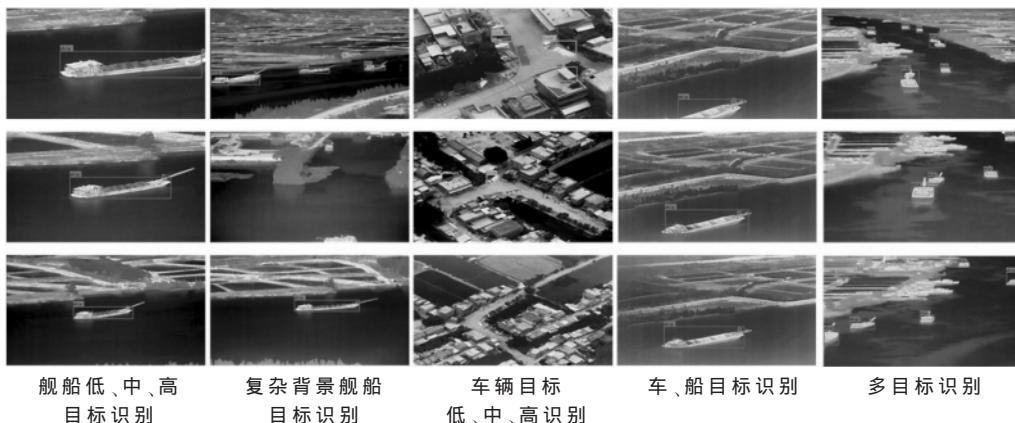


图 9 RetinaNet 识别结果

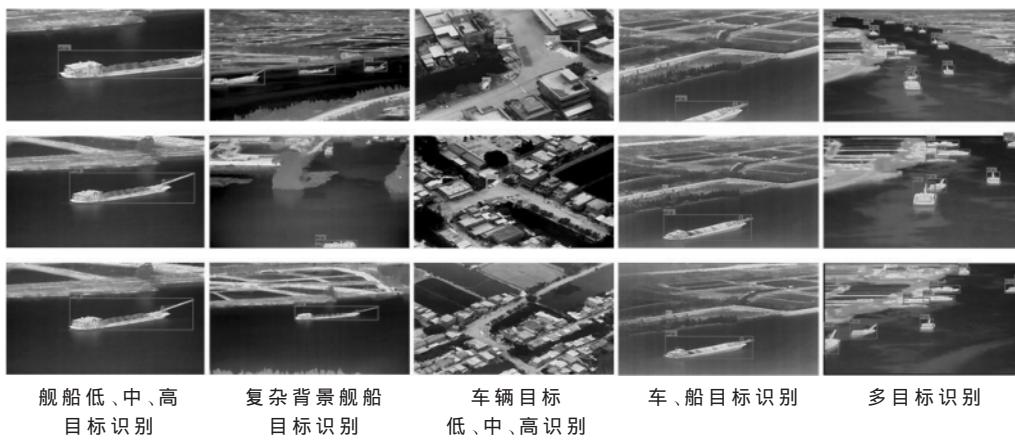


图 10 YOLOv3-SF 识别结果

表 6 红外舰船目标在不同模型上的检测结果

模型	主干网	准确率/%	召回率/%	AP/%	AP <sup>IoU=0.5</sup> /%	AP <sup>IoU=0.75</sup> /%	帧率/(f/s)
SSD500	ResNet-101	76.83	76.88	54.28	76.34	56.92	21
YOLOv3-416	DarkNet-53	80.41	80.66	60.53	80.27	61.89	39
RetinaNet-50-500	ResNet-101	84.76	84.91	62.93	84.46	64.58	14
YOLOv3-S	DarkNet-53	89.97	89.97	68.42	89.75	70.38	35
YOLOv3-F	DarkNet-53	89.61	89.72	67.88	89.55	72.36	36
YOLOv3-SF	DarkNet-53	91.74	91.95	70.64	91.72	80.47	33

表 7 红外车辆目标在不同模型上的检测结果

模型	主干网	准确率/%	召回率/%	AP/%	AP <sup>IoU=0.5</sup> /%	AP <sup>IoU=0.75</sup> /%	帧率/(f/s)
SSD500	ResNet-101	70.31	70.31	52.36	70.28	53.43	19
YOLOv3-416	DarkNet-53	76.84	76.93	53.92	76.51	54.68	36
RetinaNet-50-500	ResNet-101	82.69	82.76	60.75	82.77	62.61	13
YOLOv3-S	DarkNet-53	86.86	86.89	64.58	86.74	66.47	31
YOLOv3-F	DarkNet-53	86.47	86.54	64.29	86.33	66.10	33
YOLOv3-SF	DarkNet-53	89.19	89.19	68.47	89.23	70.32	30

Imagery in the Visible and Infrared Spectrums 也得到较高的识别精度, 同时保证识别速度, 满足实时性的要求, 识别结果如图 11 所示。

在民用目标中可以较为简单地利用无人机对目标

进行多角度采集, 但对于军用目标可能存在使用条件受限, 无法获取多角度图像的情况。虽然识别速度会有一定提升, 但识别精度会随无人机采集图像角度数量的减少有所下降。表 9 给出了目标图像采集角度数量对 6 种

表 8 三种尺度目标在不同模型上的 mAP  
(%)

模型	small	medium	large
SSD500	59.52	72.68	81.47
YOLOv3-416	63.25	74.74	86.69
RetinaNet-50-500	67.26	78.51	88.32
YOLOv3-S	74.65	85.33	94.28
YOLOv3-F	76.41	85.72	94.99
YOLOv3-SF	78.95	87.56	96.03

不同算法 mAP 的影响。

#### 4 结论

复合翼无人机从低、中、高不同高度向下视角对海上舰船目标和复杂背景中的车辆目标俯拍得到红外图像。针对没有开源多角度下视红外图像数据集的特点,构建了自定义多角度下视目标数据集。同时,在原有单步检测主干网络 Darknet-53 的基础上进行改进,提出了一种单阶段的多角度下视空基红外目标识别算法。借鉴空间金字塔池化思想,在原有主干网络 Darknet-53 中增加 SPP 模块。同时对损失函数进行改进,将原有的交叉熵损失改为 Focal loss 损失,改进后的损失更加专注于训练难分的样本,对于易分样本起到降低损失权重的作用。利用构建数据集训练网络,对红外目标进行特征提取,确定目标的检测框并给出目标类型及检测精度。对 SSD、YOLOv3、RetinaNet 与各改进算法进行对比测试。通过几种评价指标和识别结果的对比,提出的算法在提高识别精度的基础上,同时保证了识别速度,并且在复杂背景的目标识别方面具有一定的优势,适合于多种类型的红外目标图像识别。改进的算法可移植于机载目标识别系统中,只需向地面融合处理系统传输识别后的目标信息,在传输带宽一定的条件下提高数据传输效率和增加无人机个数,有效提高了系统稳定性,为后续空基红外目标探测识别系统的多尺度多角度目标识别提供理论和实验依据。

#### 参考文献

- [1] 王旭辰, 韩煜祺, 唐林波, 等. 基于深度学习的无人机载平台多目标检测和跟踪算法研究[J]. 信号处理, 2022(1): 157–163.
- [2] 刘瑞. 基于 CNN 的航空图像目标检测算法研究[D]. 成都: 电子科技大学, 2021.
- [3] 严开忠, 马国梁, 许立松, 等. 基于改进 YOLOv3 的机载平台目标检测算法[J]. 电光与控制, 2021, 28(5): 70–74.
- [4] 朱壬泰, 胡士强. 基于四旋翼无人机平台的实时多目标检测算法[J]. 计算机仿真, 2020, 37(4): 57–61.
- [5] 周子衿. 旋翼机单目视觉图像目标检测与跟踪研究[D]. 成都: 电子科技大学, 2021.
- [6] 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 & Machine Intelligence, 2017, 39(6): 1137–1149.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision, 2016.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [9] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J/OL]. (2018-06-xx)[2021-12-09]. <https://pjreddie.com/media/files/papers/YOLOv3.pdf>.
- [10] 谢晓方, 刘厚君, 张龙杰, 等. 融合轻量级 YOLOv4 与 2KCF 算法的红外舰船目标识别[J]. 兵器装备工程学报, 2021, 42(6): 175–182.
- [11] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. arXiv preprint arXiv: 1708.02002, 2017.1, 3, 4.
- [12] 杨德振, 喻松林, 冯博, 等. 基于引导滤波的分布式多波段图像融合算法[J]. 中国电子科学研究院学报, 2021, 16(9): 877–883.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only



图 11 公开数据集识别结果

表 9 采集角度数量对不同算法 mAP 的影响

采集角度数量	SSD500	YOLOv3-416	RetinaNet-50-500	YOLOv3-S	YOLOv3-F	YOLOv3-SF	(%)
单角度采集	65.97	68.41	75.29	79.34	79.13	82.26	
双角度采集	72.16	75.32	81.45	85.72	85.24	87.63	
多角度采集	74.32	79.68	84.21	88.58	88.17	91.11	

- look once : unified, real-time object detection[C]//Proceedings of CVPR, 2015: 779–788.
- [14] 张炳力, 秦浩然, 江尚, 等. 基于 RetinaNet 及优化损失函数的夜间车辆检测方法[J]. 汽车工程, 2021, 43(8): 1195–1202.
- [15] KRASIN, T, DUERIG N, ALLDRIN V, et al. Open-images: a public dataset for large-scale multi-label and multi-class image classification[Z]. 2016.
- [16] 何东, 陈金令, 王熙. 基于改进 YOLOv3 的红外行人目标检测[J]. 中国科技论文, 2021, 16(7): 8.

(收稿日期: 2021-12-09)

**作者简介:**

刘彤(1997-), 女, 硕士研究生, 主要研究方向: 深度学习、分布式信息处理。

杨德振(1988-), 通信作者, 男, 博士研究生, 主要研究方向: 光电系统设计、弱信号提取及图像处理。

宋嘉乐(1997-), 男, 硕士研究生, 主要研究方向: 迁移学习。



扫码下载电子文档

(上接第 117 页)

- 与技术, 2021, 43(1): 76–80.
- [2] 江润东, 姚金杰, 姬娜娜, 等. Ka 波段频率源建模分析与设计[J]. 电子技术应用, 2021, 47(7): 12–16.
- [3] 谢走甜, 王志伟, 李岸舟, 等. 一种新型低杂散低相噪直接合成频率源[J]. 雷达科学与技术, 2021, 19(3): 328–331.
- [4] 刘国超. 基于 PLL 低相噪快捷频率源的研究与设计[D]. 成都: 电子科技大学, 2021.
- [5] 周叶华, 叶宝盛, 程明, 等. 一种高精度纯度的 C 频段宽带频率合成器设计[J]. 电讯技术, 2018, 58(2): 219–224.
- [6] CHENAKIN A. Frequency synthesizers: current solutions and new trend[J]. Microwave Journal, 2007, 50(5): 256–266.
- [7] 李智鹏, 刘永智, 徐铭海, 等. 基于级联式偏置锁相环的低相噪宽带频率合成器[J]. 微波学报, 2014, 30(6): 9–13.

- [8] 张大鹤, 李青平. 一种超低相位噪声宽带频率合成器的设计[J]. 无线电工程, 2016, 46(2): 58–60.

- [9] 成斌, 沈文渊, 穆晓华, 等. 一种超宽带超低相位噪声频率综合器[J]. 电波科学学报, 2021, 36(4): 532–538.

- [10] 张清. 微波全频段低功耗树状谱发生器研制[D]. 成都: 电子科技大学, 2009.

- [11] 郭鑫, 程建斌. 基于 MCM 技术的小型化频率综合器设计与实现[J]. 电子与封装, 2021, 21(12): 120303.

(收稿日期: 2021-11-14)

**作者简介:**

程建斌(1982-), 男, 硕士, 工程师, 主要研究方向: 频率合成器、微波电路与系统。

钟耀霞(1982-), 女, 硕士, 副教授, 主要研究方向: 频率合成器、电路与系统。



扫码下载电子文档

(上接第 123 页)

- <http://www.analog.com>.
- [19] BEAVERS I. JESD204 转换器内确定性延迟解密[EB/OL]. (2014-xx-xx)[2021-09-xx].<http://www.analog.com>.
- [20] 李海涛, 阮林波, 田耕. FFT 方法在 ADC 有效位测试中的应用探讨[J]. 电测与仪表, 2013, 50(10): 14–17.
- [21] TI. ADS54J60 Dual-Channel, 16-Bit, 1.0-GSPS Analog-to-Digital Converter[Z]. 2019.

(收稿日期: 2021-09-22)

**作者简介:**

李海涛(1986-), 男, 博士, 工程师, 主要研究方向: 快电子学。

李斌康(1965-), 男, 博士, 研究员, 主要研究方向: 脉冲辐射探测、快脉冲电子学和系统研制。

孙彬(1979-), 男, 硕士, 副研究员, 主要研究方向: 加速器控制。



扫码下载电子文档

## 版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所