

基于改进多级特征金字塔的陆战场目标检测算法

吴 娇 王 鹏* 乔梦雨 贺咪咪

(西安工业大学电子信息工程学院 陕西 西安 710021)

摘 要 在实际陆战场环境中,常常因为所检测的实际目标(如坦克、火炮、单兵等)存在多尺度识别率低、正负样本不均衡的难题,导致无法准确实时地实现陆战场目标检测。虽然现有的 One-stage 检测器提高了一定的速率但是不能满足准确率的要求。为此提出一种基于改进的多层次特征金字塔网络(I-MLFPN)的陆战场目标检测算法,在提高识别精度的同时提升检测速度。以目标检测框架 SSD 为基础,提出改进的多层次特征金字塔网络(I-MLFPN)来构建更有效的特征金字塔,用于检测不同尺度的对象;采用 Focal loss 损失函数,解决因正负样本不均衡带来低准确率的问题。该算法将目标检测速度提高到 24.8 帧/s,目标检测准确率提高了 7%,达到 69.3%,有效解决了检测速度与准确率兼得的问题。

关键词 深度学习 多层次特征金字塔 目标检测 通道混洗

中图分类号 TP391.9 文献标志码 A DOI: 10.3969/j.issn.1000-386x.2020.10.025

TARGET DETECTION ALGORITHM OF LAND BATTLEFIELD BASED ON IMPROVED MULTI-LEVEL FEATURE PYRAMID NETWORK

Wu Jiao Wang Peng* Qiao Mengyu He Mimi

(College of Electronics and Information Engineering, Xi'an Technological University, Xi'an 710021, Shaanxi, China)

Abstract In the actual land battlefield environment, the problems of low multi-scale recognition rate and imbalance between positive and negative samples are often accounted in the actual targets detections (such as tanks, artillery, individual soldiers and other equipment), which leads to the inaccuracy of real-time target detection in the land battlefield. Although the existing One-stage detector improves the speed, it can not meet the requirements of accuracy. Therefore, a land battlefield target detection algorithm based on improved multi-level feature pyramid network (I-MLFPN) is proposed to improve the recognition accuracy and detection speed. On the basis of target detection framework SSD, an improved multi-level feature pyramid network (I-MLFPN) is proposed to construct a more effective feature pyramid for detecting objects of different scales. Focal loss function is applied to solve the problem of low accuracy caused by the imbalance of positive and negative samples. The algorithm improves the speed of target detection to 24.8 Frames/s, and the target detection accuracy is increased by 7% to 69.3%. It effectively solves the problem of both detection speed and accuracy.

Keywords Deep learning Multi-level feature pyramid Target detection Channel shuffle

0 引 言

由于现代战争更加信息化、立体化,战场信息瞬息万变。所获得的感知情报数量急剧膨胀,给利用情报进行目标的探测与识别造成了困难,仅仅通过人工判

断无法保证情报关联的正确率和实时性^[1]。因此,在新的作战思想和作战模式下,必须进一步提高系统的识别能力,增强现代陆战场侦察与反侦察的能力和手段,以适应未来形势发展的需要。尤其在反侦察过程中,陆战场目标(如火炮、坦克、装甲车、单兵)作为陆战场中的信息感知和火力打击的终端,在陆战中有着

收稿日期:2019-07-04。国家自然科学基金项目(61671362);陕西省科技厅重点研发计划项目(2019GY-022)。吴娇,硕士生,主研领域:机器视觉、深度学习。王鹏,教授。乔梦雨,硕士生。贺咪咪,硕士生。

重要的作用^[2]。这些目标通常执行毁伤、侦察、潜伏等任务,而能否准确快速检测出敌方目标不仅决定了我方战斗力强弱,还影响着战场态势,因此指挥中心需要密切关注目标终端的检测状态,根据目标检测状态实时调整作战策略^[3]。针对陆战场反侦察系统,如何准确快速地检测出目标成为提升陆战场上感知能力所亟待解决的关键问题,同时引起学者的广泛关注^[4]。

当前目标检测领域主要有三大分支。第一种是基于候选区域的深度学习目标检测算法,也称为 Two-stage 目标检测器,例如: R-CNN 通过生成候选区域,可以保证在选取较少窗口的情况下保持较高的召回率,并提升特征鲁棒性,但训练分为多个阶段,导致训练耗时、占用磁盘空间大、速度慢等^[5]; SPP-NET 算法解决了输入图片尺度限制的问题,加快目标检测的速率,但训练仍为多个阶段,固定了卷积层,无法完全适用新任务^[6];使用最广泛的 Faster-RCNN 通过引入区域生成网络来去除低层次的候选框算法,使检测器更加高效^[7],通过 RPN 与检测网络的联合训练,可实现端到端训练,然而 Faster R-CNN 还是达不到实时的目标检测^[8]; R-FCN 将全连接子检测网络替换为完全卷积子检测网络,进一步提高了 Faster R-CNN 的检测效率,但是检测速度仍然不满足需求^[9]。第二种是基于回归方法的深度学习目标检测算法,也称为 One-stage 目标检测器。例如: YOLO 算法将目标检测转换成一个回归问题,大大加快了检测速度,但没有了候选区域机制,目标不能被非常准确定位,导致检测精度不高^[10]; SSD (Single Shot MultiBox Detector) 结合了 YOLO 中的回归思想和 Faster R-CNN 中的锚机制,使用全图各个位置的多尺度区域特征进行回归,相对地提高了速度与精度^[11]。第三种是基于残差方法的深度学习目标检测算法,使用残差学习的思想学习更深的网络,从而学习到更好的特征,以提升准确率,例如: COCO2015 竞赛中提出的 152 层的深度残差网络 ResNet,实现了高准确率的目标检测;将残差网络运用于困难样本挖掘的目标检测中,在提高精度的同时使学习到的网络模型有更强的泛化能力^[12]。

针对上述问题,本文基于多层特征金字塔网络 (MLFPN)^[13] 提出一种更有效的改进的多层次特征金字塔网络 (Improved Multi-level Feature Pyramid Network, I-MLFPN),用于检测不同尺度的物体,运用 Focal loss 损失函数解决正负样本不均衡问题,同时避免上述现有方法的局限性。

本文的主要贡献如下:

1) 在多层次特征金字塔 MLFPN 的基础上,提出

了运用通道混洗模块 ShuffleNet^[14] 代替 MLFPN 网络模型中的简化 U 型模块 (TUM),用于生成一组多尺度特征,然后交替连接 ShuffleNet 和 FFMv2s 提取多级多尺度特征。运用尺度的特征聚合模块 SFAM 将特征聚合为多级特征金字塔,以此提高目标准确率与检测速率。

2) 使用一种新的 Focal loss 损失函数进行训练^[15],解决正负样本不均衡的问题,最后用非最大抑制 (NMS) 操作以产生最终结果从而提高精度^[16]。

1 相关工作

实例之间的尺度差异是对象检测任务的主要挑战之一,通常有两种策略来解决由此挑战引起的问题^[17]。第一种是检测图像金字塔中的对象,如 SNIP^[18],这种方法只能在测试时使用,大大增加内存和计算复杂性,效率急剧下降。第二种是检测从输入图像中提取的特征金字塔中的对象,如 FPN^[19],可以在训练和测试阶段进行利用。与使用图像金字塔的第一种解决方案相比,第二种方案需要更少的内存和计算成本。此外,特征金字塔构建模块可以很容易地嵌入到基于神经网络的最先进检测器中,从而产生端到端的解决方案。

尽管具有特征金字塔的物体检测器取得了不错的结果^[20],但由于它们只是根据内在多尺度金字塔结构的骨架网络构建了特征金字塔,这种骨架网络实际上是为物体分类任务设计的,因此这些方法仍然存在以下两个局限^[21]: 金字塔中的特征图对于对象检测任务而言不够典型,它们只是简单地从为对象分类任务设计的骨干网络的特征中构造;金字塔中的每个特征图主要或甚至仅从骨干网络的单层构建,它仅包含单层信息。通常,较深层中的高级特征对分类任务更具区别性,而较浅层中的低级特征可有助于对象位置回归子任务。此外,低级特征更适用于表征具有简单外观的对象,而高级特征适用于具有复杂外观的对象^[22]。实际上,在相同尺寸(例如 460 × 460)的图像中,实际对象的外观可能完全不同。例如,陆战场中的坦克、火炮和单兵可能具有相同的尺寸,但是单兵的外观要复杂得多(如单兵五官、肢体姿态等),其特征更难以提取。因此,金字塔中的每个特征图(用于检测特定尺寸范围内的对象)主要或仅由单级特征组成将致使检测性能欠佳。

本文提出的改进多层次特征金字塔网络模型 (I-MLFPN) 将网络中的不同深度的特征信息进行提取并融合,不仅解决了小目标特征信息遗漏问题,也很大

程度上提高了目标检测的识别精度,通过运用 Focal loss 函数进一步提高目标的识别率。

2 改进多层次特征金字塔网络

2.1 多层次特征金字塔模型

本文模型架构如图 1 所示。

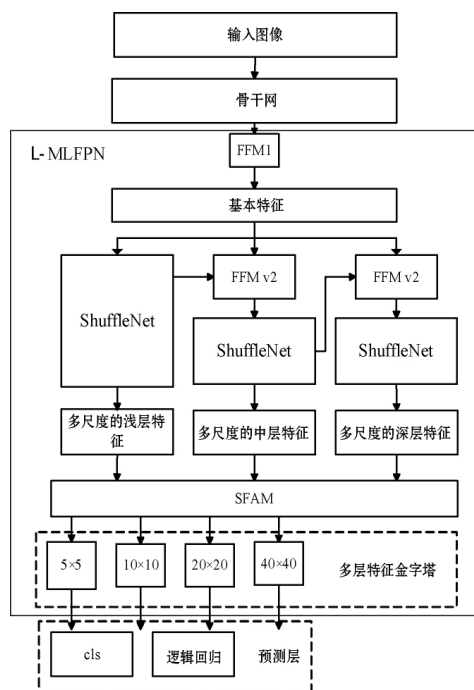


图 1 模型网络结构图

本文在 MLFPN 的基础上,提出了运用通道混洗 ShuffleNet 模块代替 MLFPN 网络模型中的 TUM 模块,优化了网络的检测精度与速度。使用骨干网和改进的多级特征金字塔网络(I-MLFPN)从输入图像中提取特征,其中 I-MLFPN 由三个模块组成,即特征融合模块(FFM)、通道混洗 ShuffleNet 模块和基于尺度的特征聚合模块(SFAM);通过 FFMv1 融合骨干网络的特征图,将语义信息丰富为基本特征;将基本信息输入给 ShuffleNet 和 FFMv2,FFMv2 融合了基本特征和前一个 ShuffleNet 的最大输出特征图,将每个 ShuffleNet 提取出的具有多级多尺度的若干特征图送入 SFAM 模块,SFAM 通过按比例缩放的特征连接操作和通道注意力机制将特征聚合到多级特征金字塔中,得出不同尺度的图像特征;根据学习的特征生成密集的边界框和类别分数;通过 Focal loss 损失函数进行训练,用非最大抑制(NMS)操作以产生最终结果。

2.2 特征融合模块 FFM

FFM 融合了图像中不同层次的特征,这对于构建最终的多级特征金字塔至关重要。它们使用 1×1 卷积层来压缩输入特征的通道,并使用连接操作来聚合

这些特征图。特别地,由于 FFMv1 以骨干网络中不同比例的两个特征图作为输入,因此它采用一个上采样操作,在连接操作之前将深度特征重新缩放到相同的尺度。FFMv1 融合浅和深的特征以产生基本特征。同时,FFMv2 采用基本特征和前一个 ShuffleNet 的最大输出特征图融合,融合后的最大特征图作为输入,并产生下一个 ShuffleNet 的融合特征。FFMv1 和 FFMv2 的结构细节如图 2 所示。

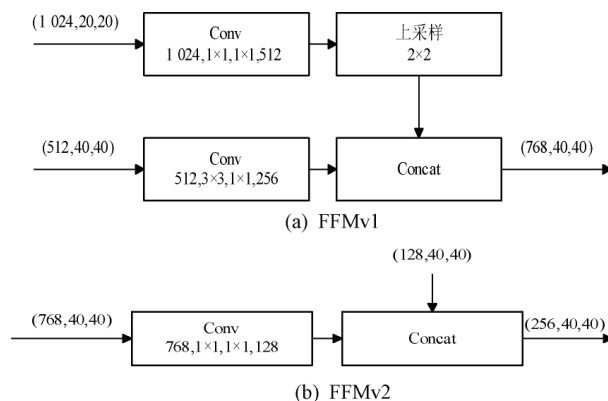


图 2 FFM 结构

2.3 通道混洗 ShuffleNet 模块

ShuffleNet 模块用于对不同深度层次的特征信息进行提取,生成多级多尺度特征。该模块主要进行两个操作:逐点群卷积和通道混洗(channel shuffle),与现有先进模型相比在类似的精度下大大降低了计算量。逐点群卷积帮助降低计算复杂度,但昂贵的逐点卷积造成有限的通道之间充满约束,这会显著地损失精度,降低通道组之间的信息流通和信息表示能力,故提出通道混洗帮助信息流通。基于这两种技术构建的高效架构为 ShuffleNet,相比于其他先进模型,对于给定的计算复杂度预算,ShuffleNet 允许使用更多的特征映射通道,有助于编码更多信息。

若允许组卷积能够得到不同组的输入数据,那么输入和输出通道会是全关联的。具体来说,对于上一层输出的通道,可做一个混洗(Shuffle)操作(如图 3 所示),再分成几个组,提供给下一层,ShuffleNet 可以允许使用更多的特征映射通道,这样就可以编码更多信息。

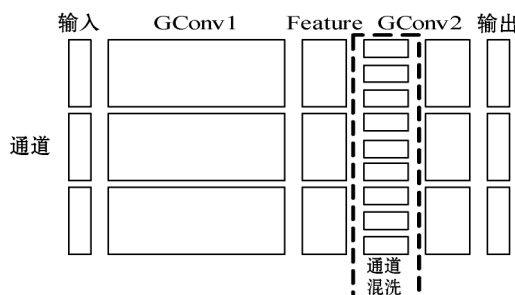


图 3 通道混洗结构图

图4 为构建的一个 ShuffleNet 单元,便于实现实际模型。

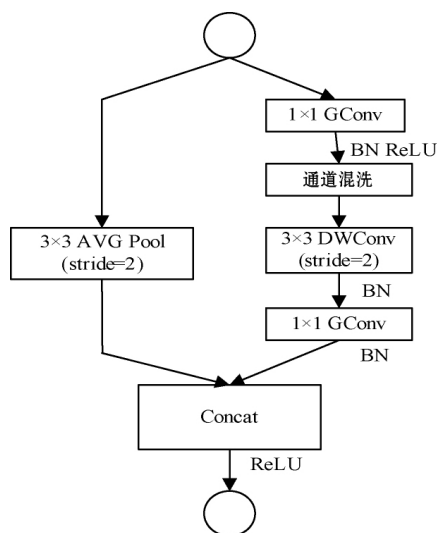


图4 混洗单元

图4 是做降采样的 ShuffleNet 单元,主要做了两点修改:(1) 在辅分支加入步长为2的 3×3 平均池化;(2) 原本做元素相加的操作转化为了通道级联,扩大了通道维度,增加的计算成本却很少。

ShuffleNet 单元可以高效地计算,归功于逐点卷积和通道混洗。相比于其他先进单元,在相同设置下复杂度降低,网络模型速率得到提升。

第一个 ShuffleNet 没有任何其他 ShuffleNet 的先验知识,因此它只从 X_{base} 学习。输出多级多尺度特征的计算如下:

$$[X_1^l, X_2^l, \dots, X_i^l] = \begin{cases} T_l(X_{base}) & l=1 \\ T_l(F(X_{base}, X_i^{l-1})) & l=2, 3, \dots, L \end{cases} \quad (1)$$

式中: X_{base} 表示基本特征; x_i^l 表示在第 l 个 TUM 中具有第 i 个尺度的特征; L 表示 TUM 的数量; T_l 表示第 l 个 TUM 处理; F 表示 FFMv1 处理。

2.4 特征聚合模块 SFAM

SFAM 旨在将由 ShuffleNet 模块生成的多级多尺度特征聚合成多级特征金字塔。SFAM 的第一阶段是沿着信道维度将等效尺度的特征连接在一起。聚合特征金字塔可以表示为 $X = [X_1, X_2, \dots, X_i]$, 其中 $X_i = \text{Concat}(x_i^1, x_i^2, \dots, x_i^L) \in \mathbf{R}^{w_i \times h_i \times C}$, 指的是尺度第 i 个最大的特征。这里, 聚合金字塔中的每个比例都包含来自多级深度的特征。在第二阶段, 引入了通道注意模块, 以促使特征集中在最有益的通道。在 Squeeze-and-Excitation Block(SE Block) 之后, 使用全局平均池化在挤压步骤中生成通道统计 $z \in \mathbf{R}^C$ 。为了完全捕获通道依赖性, 以下激励步骤通过两个全连接层学习注意机制, 如图5所示。

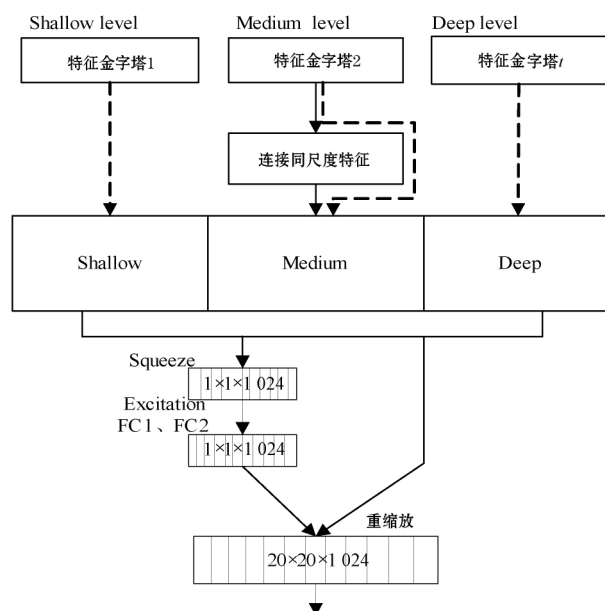


图5 按缩放比例的特征聚合模块

SFAM 的第一阶段是沿着信道维度连接具有等效比例的特征, 并得到对输入 z 重新加权的特征聚合 F_{ex} , 并表示为 s ; 第二阶段使用 SE 注意机制以自适应方式聚合特征, 对于第 i 个特征通道数为 c 的特征层 X_i^c 引入注意力模块 s_c 进行加权聚合操作, 构成聚合特征层 F_{scale} , 表示为 \tilde{X}_i^c , 公式如下:

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

$$\tilde{X}_i^c = F_{scale}(X_i^c, s_c) = s_c \cdot X_i^c \quad (3)$$

式中: σ 表示 ReLU 函数; δ 表示 Sigmoid 函数; 权值参数 $W_1 = R_1^{r \times C}$, $W_2 = R_2^{r \times C}$, 其中 r 是减少的比例, 本实验设置为 $r_1 = 16$, $r_2 = 8$; C 为特征通道数; R 表示每个特征增强或是减弱的随机因子数。

2.5 Focal loss 函数

针对一阶段检测器提出 Focal Loss 函数, 训练过程中有前景和背景类别之间的完全不平衡存在, 这会主导交叉熵损失。难分负样本在梯度和损失中占据主导地位, 将损失函数变形降低易分样本的权重, 专注于训练难分负样本。Focal loss 加了 $(1 - m_i)^\beta$ 到交叉熵上, β 是可以调节的专注参数 $\beta > 0$, m_i 是不同类别的分类概率, 则 Focal loss 定义为:

$$FL(m_i) = -(1 - m_i)^\beta \log(m_i) \quad (4)$$

在实际中, 可以运用添加一个权重 w_i 平衡的变体的 Focal loss 函数:

$$FL(m_i) = -w_i (1 - m_i)^\beta \log(m_i) \quad (5)$$

Focal loss 通过调整 loss 的计算公式使单级结构达到和 Faster RCNN 一样的准确度。 $1 - m_i$ 为调制因子, 当 m_i 很小时, 那么调制因子 $1 - m_i$ 接近 1, 损失不受影响; 当 $m_i \rightarrow 1$, 因子 $1 - m_i$ 接近 0, 那么分的比较好的样

本的权值就被调低。专注参数 β 平滑地调节了易分样本调低权值的比例, β 增大能增强调制因子的影响。

3 实 验

3.1 实验环境

本文算法在 Ubuntu 16.04 操作系统, 采用深度学习框架 Pytorch, 实现了基于多层次特征金字塔的陆战场目标检测算法。实验平台采用 CPU: Intel(R) Core(TM) i5-8600 3.10 GHz; 内存: 16 GB; GPU: NVIDIA GTX 1080Ti。为了验证算法的准确率和实时性, 选取了当前性能较好的几种算法在同样的环境下进行对比测试。训练集是随机从数据集中抽取的 70% 的数据, 测试集是从数据集中抽取的 30% 的数据。

3.2 实验数据集

考虑到实际陆战场中所包含的目标种类, 本文所要识别的目标种类包含坦克(Tank)、坦克炮塔(Cannon)、装甲车辆(Car)、人(Person)和枪(Gun)。

由于目标种类特殊, 公开资源较少。因此数据集一部分来自于内部资料, 包含上述所提目标的图片 9 000 幅。数据通过添加噪声, 以及一定程度上的尺度变换进行扩充后达到 27 000 幅。为了使数据更多样化, 自行从网络上寻找包含目标的相关视频图像制作数据集, 采集了 3 000 幅。

经过扩充的原始资料以及网络上的图像, 数据集一共有 30 000 幅图片。每一幅图片按照 PASCAL VOC 数据集的格式进行了人工标注, 数据集部分图片如图 6 所示。



图 6 部分数据集

3.3 损失函数对比

本文将 Focal loss 作为模型的损失函数。为了证明使用 Focal loss 的优越性, 将本文所提目标识别算法的其他部分保持不变, 使用深度学习中其他常用的损

失函数 Cross entry 和 Center loss, 在同样环境下进行训练和测试。

(1) Cross Entry 损失函数:

$$H(p, q) = - \sum_{i=1}^N p(x^{(i)}) \log q(x^{(-i)}) \quad (6)$$

交叉熵是用来评估当前训练得到的概率分布与真实分布的差异情况, 减少交叉熵损失就是在提高模型的预测准确率。式中: $p(x)$ 是指真实分布的概率; $q(x)$ 是模型通过数据计算出来的概率估计。

(2) Center loss 损失函数:

$$CL = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (7)$$

式中: c_{y_i} 表示第 y_i 个类别的特征中心; x_i 表示全连接层之前的特征; m 表示小批梯度下降的大小。因此式(7)就是希望一批中的每个样本的特征离特征的中心距离的平方和要越小越好, 也就是类内距离要越小越好。

ROC 曲线又称接收者操作特性曲线, 用来评价分类器的性能。横坐标: 伪正类率(False Positive Rate, FPR), 预测为正但实际为负的样本占有所有负例样本的比例; 纵坐标: 真正类率(True Positive Rate, TPR), 预测为正且实际为正的样本占有所有正例样本的比例, 运用 ROC 曲线绘制三种损失函数, 如图 7 所示。

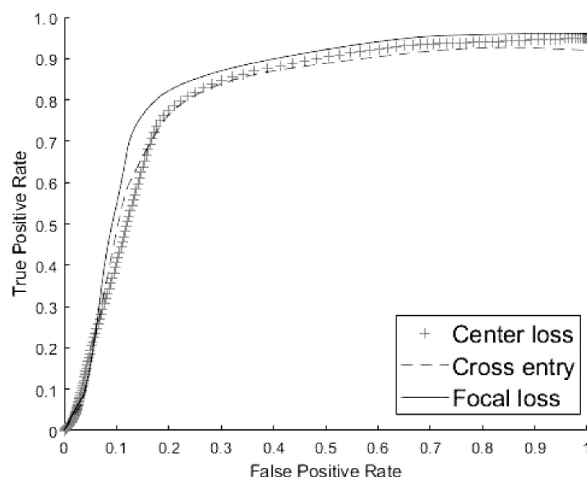


图 7 ROC 测试曲线图

可以看出, 在二分类模型中, Focal loss 损失函数性能更好, 使得分类器效果最优。

3.4 算法对比

为了验证算法的识别准确率, 将本文算法模型与当前主流的目标检测模型对比, 所对比的性能参数由 FPS(每秒帧率)、AP(平均正确率)、AP50 和 AP75(以

0.5 和 0.75 为 IoU 临界值分别估计出平均准确度), 结果如表 1 所示。

表 1 不同算法精度比较表

方法	骨干网络	FPS	AP/%	AP50/%	AP75/%
方法 1	—	9.5	33.2	53.4	35.7
方法 2	ResNet-101	12.4	39.1	62.9	44.6
方法 3	—	7.1	42.8	61.1	45.1
方法 4	ResNet-101	24.8	45.6	69.3	50.2
方法 5	VGG-16	25.1	39.2	60.9	44.3

表 1 中: 方法 1 为 DSSD, 添加了反卷积层融合多层特征的 SSD 网络; 方法 2 为 RefineDet, 引入类似 FPN 网络的特征融合操作用于检测网络, 检测框架还是 SSD; 方法 3 为 Cascade R-CNN, 采用级联结构的 RCNN 网络; 方法 4 为本文算法, 其中骨干网络采用 ResNet-101; 方法 5 为本文算法, 其中骨干网络采用 VGG-16。

为了验证本文模型的速度, 实验环境均采用 GPU 加速, 读取 20 000 幅图片, 图片大小为 460×460 。采用 FPS(每秒帧率)作为评价标准, 一帧即读取一幅图片的时间, FPS 就是一秒读取的帧数。本文采用的 ResNet-101 网络模型的 FPS 可以达到 24.8, 优于采用同样骨干网络的 DSSD、RefineDet 以及 Cascade R-CNN 算法。同时, 保持其他网络参数与结构不变, 采用 VGG-16 作为骨干网络进行实验, 模型参数更少, 相对于采用 ResNet-101 作为骨干网络的方法, 其检测速度提升了 1%, FPS 达到 25.1。

由表 1 可知, DSSD 算法的平均正确率低于本文算法, 由于通道混洗模块减少参数, 本文算法正确率高出 12.4%; RefineDet 平均正确率接近于采用 VGG-16 作为特征提取网络的本文算法; Cascade R-CNN 方法平均正确率略高于采用 VGG-16 作为特征提取网络的本文方法, 但检测速度更慢; 采用 ResNet-101 作为骨干特征提取网络的本文方法在平均正确率上优于使用 VGG-16 方法, 检测精度提升了 6%, 检测速度仅低了 0.3 帧/s。在每秒帧率与平均正确率上, 本文算法均优于现有其他主流目标识别算法。综合识别准确率与目标检测速度的情况下, 使用 ResNet-101 作为骨干提取网络优于使用 VGG-16 的方法。部分测试结果如图 8 所示。



图 8 部分测试结果图

4 结 语

本文提出了基于多层次特征金字塔的陆战场目标检测算法。将骨干网提取的多层特征融合为基本特征, 然后将其馈入交替连接的 ShuffleNet 模块和特征融合模块, 以提取更具代表性的多层次多尺度特征。显然, 这些 ShuffleNet 层比主干中的层深得多, 因此用于物体检测时可以表现出更多的特征。运用 Focal loss 函数可以较好地处理正负样本不均衡的问题。与现有其他检测算法相比, 本文算法目标检测准确率更优, 但本文的整体网络模型架构较复杂, 如果对模型进行修建, 未来可以实现嵌入式终端的移植。

参 考 文 献

- [1] 张晓海, 操新文, 耿松涛, 等. 基于深度学习的军事辅助决策智能化研究[J]. 兵器装备工程学报, 2018, 39(10): 162-167.
- [2] 张敬慧, 胡毅钧. 战场目标特性与燃烧弹药结构特性分析[J]. 科学技术创新, 2018(2): 21-22.
- [3] 王宝峰, 李战武, 寇英信, 等. 空战对敌目标攻击路径辅助控制研究[J]. 计算机仿真, 2017, 34(12): 55-60.
- [4] 张桂林, 王菁, 王惠娟, 等. 基于知识图谱的战场目标智能识别推理研究[C]//第六届中国指挥控制大会论文集(上册). 北京: 电子工业出版社, 2018.
- [5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2014: 580-587.
- [6] Purkait P, Zhao C, Zach C. SPP-Net: Deep absolute pose regression with synthetic views[EB]. arXiv: 1712.03452,

2017.

- [7] Ren S Q, He K M, 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.
- [8] Zhong Z Y, Sun L, Huo Q. An anchor-free region proposal network for faster R-CNN based text detection approaches[EB]. arXiv: 1804.09003 2018.
- [9] Dai J, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks[EB]. arXiv: 1605.06409 2016.
- [10] Heimer R, Myrseth K O R, Schoenle R. YOLO: Mortality beliefs and household finance puzzles[J]. Journal of Finance, 2019 74(6): 2957–2996.
- [11] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//2016 European Conference on Computer Vision. Springer 2016: 21–37.
- [12] 张超 陈莹. 残差网络下基于困难样本挖掘的目标检测[J]. 激光与光电子学进展 2018 55(10): 105–111.
- [13] Zhao Q, Sheng T, Wang Y, et al. M2Det: A single-shot object detector based on multi-level feature pyramid network[C]//The 33rd AAAI Conference on Artificial Intelligence 2019.
- [14] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[EB]. arXiv: 1707.01083 2017.
- [15] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2020 42(2): 318–327.
- [16] Neubeck A, Gool L V. Efficient non-maximum suppression[C]//The 18th International Conference on Pattern Recognition. IEEE 2006: 850–855.
- [17] Shan Q W, Zheng X B, He X H, et al. Fast object detection and recognition algorithm based on improved multi-scale feature maps[J]. Laser and Optoelectronics Progress 2019 56(2): 021002.
- [18] Singh B, Davis L S. An analysis of scale invariance in object detection-SNIP[EB]. arXiv: 1711.08189 2017.
- [19] Dubnicki C, Ungureanu C, Kilian W. FPN: A distributed Hash table for commercial applications[C]//IEEE International Symposium on High Performance Distributed Computing, 2004.
- [20] 陈景明 金杰 王伟锋. 基于特征金字塔网络的改进算法[J]. 激光与光电子学进展 2019 56(21): 211505.
- [21] 任之俊 蔺素珍 李大威, 等. 基于改进特征金字塔的 Mask R-CNN 目标检测方法[J]. 激光与光电子学进展, 2019 56(4): 041502.
- [22] Dollar P, Appel R, Belongie S, et al. Fast feature pyramids for object detection[J]. IEEE transactions on pattern analysis and machine intelligence 2014 36(8): 1532–1545.

(上接第 137 页)

- [10] Zhong Z, Fan B, Duan J, et al. Discriminant tensor spectral-spatial feature extraction for hyperspectral image classification[J]. IEEE Geoscience Remote Sensing Letters, 2015, 12(5): 1028–1032.
- [11] Tao M, Zhou F, Liu Y, et al. Tensorial independent component analysis-based feature extraction for polarimetric SAR data classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53(5): 2481–2495.
- [12] Ren Y, Liao L, Maybank S J, et al. Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis[J]. IEEE Geoscience & Remote Sensing Letters, 2017, 14(9): 1431–1435.
- [13] Kilmer M E, Martin C D. Factorization strategies for third-order tensors[J]. Linear Algebra and its Applications 2011, 435(3): 641–658.
- [14] Zhang Z, Ely G, Aeron S, et al. Novel methods for multilinear data completion and de-noising based on tensor-SVD[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [15] 许丽 彭尧. 一种改进的奇异值分解脑电去噪方法[J]. 信息技术 2018(5): 149–152, 159.
- [16] 胡卫红 舒泓 栾宇光. 基于奇异值分解的电能质量信号去噪[J]. 电力系统保护与控制 2010 38(2): 30–33.

(上接第 154 页)

- [5] Chen L, Wu Y, Dsouza A M, et al. MRI tumor segmentation with densely connected 3D CNN[C]//Houston: International Conference on Medical Imaging and Image Processing 2018.
- [6] 王刚 陈永光 杨锁昌, 等. 采用图像块对比特性的红外弱小目标检测[J]. 光学精密工程, 2015 23(5): 1424–1433.
- [7] Jiang B, Zhang L, Lu H, et al. Saliency detection via absorbing markov chain[C]//2013 IEEE International Conference on Computer Vision 2013.
- [8] 侯庆岑 潘晨 杨勇. 基于视觉显著度的多聚焦图像融合方法[J]. 计算机工程 2015 41(1): 211–217.
- [9] 朱婷 王瑜 肖洪兵, 等. 基于 WRN-PPNet 的多模态 MRI 脑肿瘤全自动分割[J]. 计算机工程 2018 44(12): 258–263, 270.
- [10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention 2015.
- [11] Kistler M, Bonaretti S, Pfahrer M, et al. The virtual skeleton database: an open access repository for biomedical research and collaboration[J]. Journal of Medical Internet Research, 2013 15(11): 245.