

doi: 10.7690/bgzd.2020.07.006

## 卷积神经网络在炮兵对抗训练系统中的应用

陈 栋<sup>1</sup>, 杨传栋<sup>1</sup>, 秦 杰<sup>1</sup>, 蒋滨安<sup>1</sup>, 修德良<sup>2</sup>

(1. 陆军炮兵防空兵学院高过载弹药制导控制与信息感知实验室, 合肥 230031;

2. 中国人民解放军 77611 部队 53 分队, 拉萨 850000)

**摘要:** 针对炮兵对抗训练系统中炸点图像目标捕捉的问题, 提出一种基于 YOLACT 的炸点区域快速识别及分割方法。对特征提取网络结构和参数进行修改, 结合预测分支网络和掩膜生成网络输出炸点位置和区域范围, 根据区域信息得到炸点中心坐标。实验结果表明: 在构建的炸点数据集上, 该方法能准确地识别和分割炸点目标, 速度达到 21.2 fps, 整体上优于对比算法, 能较好地解决炮兵对抗训练系统中的问题。

**关键词:** 对抗训练; 炸点识别; 卷积神经网络; 掩膜生成

**中图分类号:** TP391 **文献标志码:** A

## Application of CNN in Artillery Countermeasure Training System

Chen Dong<sup>1</sup>, Yang Chuandong<sup>1</sup>, Qin Jie<sup>1</sup>, Jiang Bin'an<sup>1</sup>, Xiu Deliang<sup>2</sup>

(1. Laboratory of Guidance Control &amp; Information Perception Technology of High Overload Projectiles,

PLA Army Academy of Artillery &amp; Air Defense, Hefei 230031, China;

2. No. 53 Team, No. 77611 Unit of PLA, Lasa 850000, China)

**Abstract:** A fast burst point area identification and segmentation algorithm based on YOLACT is proposed to capture the blast point in artillery countermeasure training system. Firstly, the feature extraction network structure and parameters are modified for the target of the blast point area. The prediction branch network and the mask generation network are combined to output the location and boundary area of the blast point. Finally, the location of the blast point is calculated according to the boundary information. The experimental results show that the method in this paper can accurately identify and segment the target of the blast point on the constructed blast point data set, and the speed reaches 21.2 fps, which is better than the comparison algorithm as a whole, and can solve a basic problem in the artillery confrontation training system.

**Keywords:** countermeasure training; blast point identification; convolutional neural network (CNN); mask generation

### 0 引言

炮兵对抗训练中, 炸点快速准确识别是导调评估的必要前提, 对系统有序运行至关重要。当前炸点识别方法通常包括声测<sup>[1]</sup>、激光扫描<sup>[2]</sup>、雷达探测<sup>[3]</sup>和图像目标识别<sup>[4]</sup>4种: 声测法是借助若干声波传感器接收的炸点声波信号进行结算, 实现空间定位, 测试精度较低, 抗噪能力差; 激光扫描利用主动式激光在靶场上空高速扫描形成扫面平面, 捕捉经过平面的炸点, 此类方法耗费巨大; 雷达探测法利用雷达定向发射的脉冲电波, 通过目标回波探测其方位, 可测量弹道, 但系统庞大, 成本较高; 图像法利用多台相机捕捉地面炸点, 设备成本低, 识别速度快, 但由于炸点目标多, 容易漏检误检。

近年来, 随着核心算法的突破、计算能力的提高以及海量数据的支撑, 基于卷积神经网络(convolutional neural network, CNN)的目标识别算法取得了超过传统算法的良好表现, 在军事领域中也

得到快速拓展和应用。目前该方法可分为基于候选框的两阶段识别算法<sup>[5]</sup>、基于回归的单阶段识别算法<sup>[6]</sup>及无候选框的单阶段识别算法<sup>[7]</sup>, 其中单阶段识别算法在识别速度上达到了实时的要求。炸点目标由于烟尘或火光形状多样, 仅仅使用矩形框进行识别往往并不能提供精确的定位。文献[8]在矩形检测框的基础上实现目标分割, 能够更准确地指示目标。为了快速准确识别炸点, 笔者提出一种基于全卷积神经网络(you only look at coefficients, YOLACT<sup>[9]</sup>)的炸点区域快速识别及分割方法。

### 1 对抗训练系统

实兵演练对抗中, 弹药爆炸位置不是真正落在演训双方装备上; 因此, 弹药作战效能和战术运用效果难以真实体现<sup>[10]</sup>。针对对抗式训练现存问题, 构建了一套炮兵对抗训练系统, 通过虚拟现实技术将实际目标进行坐标映射到真实炸点位置, 导调部

收稿日期: 2020-03-19; 修回日期: 2020-04-27

基金项目: 军队“十三五”装备预研项目(301070103)

作者简介: 陈 栋(1981—), 男, 安徽人, 博士, 副教授, 从事智能弹药总体设计、炮兵作战研究。E-mail: 1092850034@qq.com。

利用显示设备和评估系统监测、导调和评估目标毁伤。

基于虚拟现实的炮兵对抗训练系统包括战场监测系统、导调评估系统和坐标映射系统。战场监测系统利用高台摄像机对炸点进行捕捉，并解算炸点坐标；导调评估系统用来串联对抗单元、获取对抗数据，制定对抗规则，评估对抗效能；坐标映射系统用来构建数字地图，通过增强现实技术投影映射坐标信息，实时成像。

阵地区域布设如图 1 所示。将训练场地划分为 4 块区域，红蓝双方观察所人员、炮阵地完全按照真实战场环境作业，在射击时与真实射向基准旋转  $90^\circ$ 。

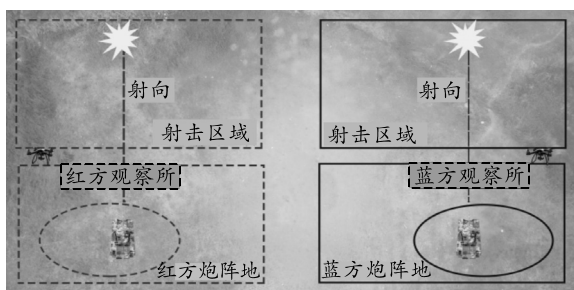


图 1 系统阵地设置

由于基准射向的旋转，炸点落于射击区而未落入敌方阵地，而是通过导演部的导调评估系统，将虚拟目标信息采用增强现实技术以可视化形式投影在双方观察所人员视场中，观察所人员依据视场中实际炸点信息和虚拟目标的位置关系进行校正射击；导演部同时结算炸点坐标和双方阵地人员装备之间的位置关系，根据毁伤标准和退出机制将结果实时传递给参训人员，直到对抗结束。

## 2 基于 YOLACT 的炸点目标识别算法

### 2.1 目标识别模型

笔者使用基于 YOLACT 的目标识别算法对炸点区域进行识别和分割。整个算法框架如图 2 所示，输入待检测图像或视频，利用特征提取网络在不同尺度上提取图像特征，分别送入预测分支和掩膜分支网络；掩膜生成网络得到一系列基本掩膜模板，预测分支网络对目标的类别、位置和模板系数矩阵进行预测；通过非极大值抑制和得分阈值筛选后，将基本掩膜模板和模板系数矩阵进行线性组合得到最终目标掩膜。最后输出预测类别、边界框位置和对应的炸点分割结果。

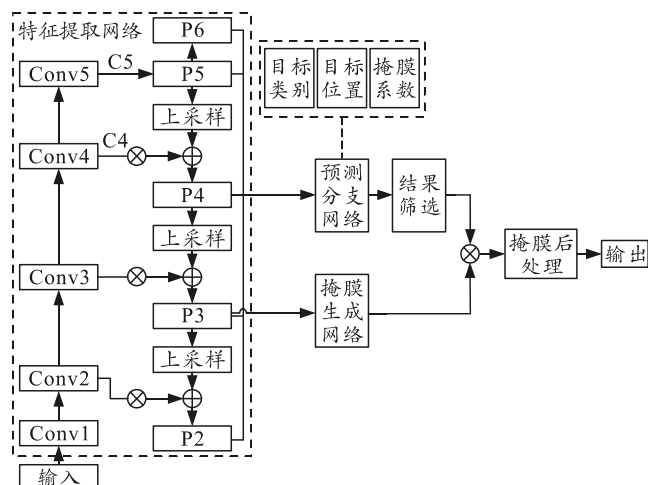


图 2 炸点识别算法框架

### 2.2 特征提取网络

在一个多层卷积神经网络中，低层特征通常能更好地表征图像纹理、边缘等细节信息，经过多层卷积操作后，神经元感受野扩大，高层特征能够更好地表征语义信息，但相应会忽略图像细节信息。对于炸点目标，高层语义信息较为简单，并不需要神经元具有很大感受野，而是更多依靠低层细节信息对炸点进行识别。特征金字塔结构 (feature pyramid network, FPN)<sup>[11]</sup>能够对低层语义信息充分利用，将其与高层信息相融合，结合在不同尺度的预测结果，能达到更好的识别效果。本模型选择了 Resnet50<sup>[12]</sup>+FPN 作为特征提取网络，网络基本结构如图 2 左侧所示。其中 Conv\_x 为 Resnet50 的残差块结构，通过跳层连接方式提高了特征提取能力。P\_x 作为特征提取网络的在不同尺度上的输出，是后续预测和掩膜网络使用的特征图。经过特征提取网络前向传播得到各层残差块结构的输出特征 {C2, C3, C4, C5}。随后利用 FPN 融合方式，将各层特征与对应的经过上采样的高层特征进行相加，最终得到融合特征 {P2, P3, P4, P5, P6}。由于炸点尺度较小并考虑时间消耗，经过对比实验，最终选择去除 P2 和 P6 层特征，各层对应的锚点相对于输入图像的尺寸范围选择 { $16^2, 32^2, 64^2$ }，每个尺寸对应  $a$  种尺度的锚点。为了提高算法针对性，对训练集目标数据的尺度进行 K-means 聚类， $a=3$  时，即尺度比设为 {1:1.3, 1:1, 1.3:1} 3 种，可以达到更优的训练准确度。

### 2.3 掩膜模板生成

#### 2.3.1 掩膜生成网络

掩膜生成网络是一个全卷积网络，参数少，相

对于全连接网络能够保留完整空间信息。网络以 P3 层特征为输入产生  $k$  层掩膜，依靠专门的分类分支

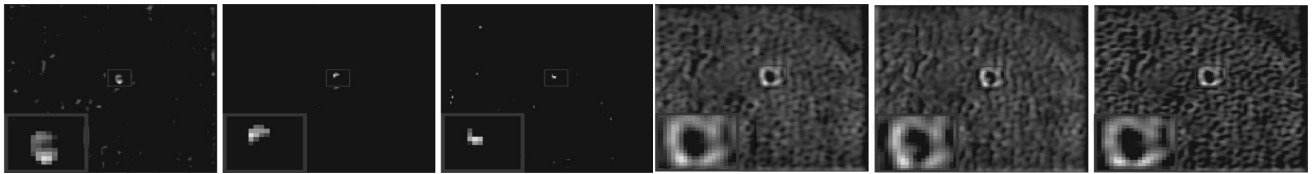
进行预测类别，并凭此选择输出对应的掩膜。具体结构设计如表 1 所示。

表 1 掩膜生成网络结构

类型	输入(P3)	卷积层(3×3)	卷积层(3×3)	卷积层(3×3)	反卷积层(3×3)	卷积层(3×3)	卷积层(1×1)+ReLU
输出尺寸	69×69×256	69×69×256	69×69×256	69×69×256	138×138×256	138×138×256	138×138×k

首先对固定尺寸的特征图(69×69×512)进行 3 次相同的卷积操作，然后利用反卷积操作增加尺度(138×138×256)，最后利用 1\*1 卷积并通过修正线性单元 ReLU 进行修正，得到分辨率是 138×138、维度为  $k$  的掩膜图。得到的掩膜模板输出结果如图 3

所示。从图中可以看到：每个掩膜对应不同的属性，如掩膜(a)(b)(c)对炸点的不同位置进行了激活，分别对应下、上和中部部位，并对背景部分抑制明显，提取出了作为前景的炸点区域；掩膜(d)(e)(f)对区域边界进行了激活，在不同的方位各有所侧重。



(a) 激活下部掩膜 (b) 激活上部掩膜 (c) 激活中部掩膜 (d) 激活炸点边界 (e) 激活左侧边界 (f) 激活下方边界

图 3 基本掩膜输出示意

### 2.3.2 预测分支网络

在得到诸多掩膜模板后，为了输出目标掩膜，笔者增加了掩膜系数预测分支如表 2 所示。

表 2 预测分支网络结构

类型	输入(Pi)	卷积层(3×3)	卷积层(3×3)+ReLU
输出尺寸	$W \times H \times 256$	$W \times H \times 256$	$W \times H \times 256$
类型	分类分支(3×3)	位置分支(3×3)	掩膜系数分支(3×3)+tanh
输出尺寸	$W \times H \times a \times c$	$W \times H \times a \times 4$	$W \times H \times a \times k$

预测分支在 {P3,P4,P5} 各层进行预测，每层尺度为  $W \times H$ ，首先经过 2 层卷积层及 ReLU 层，然后并行输出分类、位置和掩膜系数矩阵。其中， $c$  为目标种类，此处仅有炸点目标一类，转化为背景和炸点的二分类问题。掩膜系数矩阵为  $k$  个模板对应的权重系数，作为对上节基本掩膜选择和融合的依据。为了保证网络非线性，在输入掩膜系数分支后增加 tanh 激活函数，能够使输出更加稳定。掩膜系数预测如图 4 所示。图中取  $k=32$ ，并将预测的掩膜系数按绝对值大小排序，表示每个基本掩膜对最终掩膜的贡献。

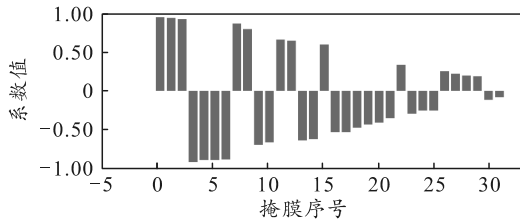


图 4 掩膜系数预测结果

由于网络对每层中各个位置都进行了预测，存在大量的无用候选结果，需要对初步预测结果进行

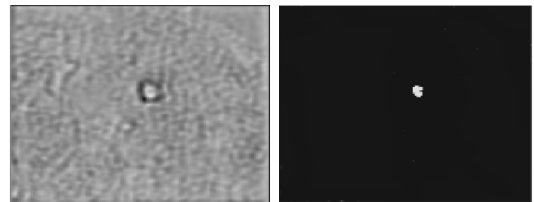
筛选。对于仅有单炸点情况，通过使用非极大值抑制和置信度阈值限制后，得到了一组最终的预测结果。

### 2.4 掩膜线性组合

利用最终掩膜系数结果对掩膜模板进行融合，得到目标实例的分割结果如图 5，可用下式表示：

$$M = \sigma(PC^T) \tag{1}$$

其中： $P$  为  $138 \times 138 \times k$  维的基本掩膜矩阵； $C$  为  $n \times k$  维的掩膜系数； $n$  为经过非极大值抑制和置信度阈值限制后的结果个数，并利用 sigmoid 激活函数  $\sigma()$  对其进行非线性约束。可以看到掩膜结果是针对整个特征图尺度的，而我们仅需要目标边界框范围；因此，在后处理阶段对掩膜结果使用预测得到的边界框进行裁剪。



(a) 掩膜线性组合 (b) 结果阈值二值化

图 5 掩膜线性组合结果

### 2.5 多任务损失函数

本方法使用 3 种损失联合训练模型：分类置信度损失  $L_{cls}$ 、边界框回归损失  $L_{box}$  和掩膜损失  $L_{mask}$ 。为了保证各个损失对总损失贡献基本持平，设置了平衡系数  $\lambda_{cls}$ 、 $\lambda_{box}$ 、 $\lambda_{mask}$  对 3 种损失进行加权。多

任务损失函数如下式：

$$\begin{aligned}
 L &= \lambda_{\text{cls}} L_{\text{cls}} + \lambda_{\text{box}} L_{\text{box}} + \lambda_{\text{mask}} L_{\text{mask}}, \\
 L_{\text{box}} &= \sum_{i \in \text{Pos}} \sum_{m \in \{c_x, c_y, w, h\}} \text{Smooth}_{L1}(l_i^m - g_j^m), \\
 L_{\text{cls}} &= - \sum_{i \in \text{Pos}} \log(p_i^{\text{bomb}}) - \sum_{i \in \text{Neg}} \log(p_i^0), \\
 L_{\text{mask}} &= \sum_{i \in \text{Pos}} \text{BCE}(M_i, M_j^{\text{gt}}). \quad (2)
 \end{aligned}$$

其中： $p_i^{\text{bomb}}$ 为第  $i$  个预测框对应炸点的概率； $p_i^0$ 为被预测为背景的概率； $i \in \text{Pos}$  为正样本，即预测框和真值框交并比大于阈值； $i \in \text{Neg}$  为负样本； $N$  为正样本个数； $l_i^m$  为第  $i$  个预测框的预测值； $g_j^m$  为第  $j$  个真值； $m$  分别为预测框的位置和长宽  $\{c_x, c_y, w, h\}$ ； $\text{Smooth}_{L1}$  为平滑的 L1 损失； $M_i$  为第  $i$  个融合后的炸点区域预测掩膜值； $M_j^{\text{gt}}$  为第  $j$  个真值； $\text{BCE}()$  为像素级的二值交叉熵。

### 3 实验

#### 3.1 数据集及参数设置

在草地、沙漠、土地等不同场景下，可利用电视经纬仪、旋翼无人机等采集设备在不同高度和距离上对各类炸点爆炸进行拍摄，构建了炸点数据集，并按照 COCO 格式制作数据集样本。同时，使用数据增强方法<sup>[6]</sup>，对总的数据进行亮度、饱和度、色调、颜色灰度扭曲等变换，获取约 2 000 个样本数据，其中 1 700 张图片用作训练集，300 张用于测试集。通过 labelme 工具包对样本图像进行标注，使其满足 COCO 数据集格式，以便进行训练。

实验硬件使用配置有内存为 16 GB 和英伟达 1070 GPU 的 PC 机；操作系统为 ubuntu16.04，在 python 中基于 pytorch 深度学习框架对算法模型编程实现。训练中部分参数设置为：学习率取 0.001，动量取 0.9，权重衰减为 0.000 5，批处理图像数为 5，平衡系数设为  $\lambda_{\text{cls}}=1$ ， $\lambda_{\text{box}}=1.5$ ， $\lambda_{\text{mask}}=6.25$ ，检测非极大值抑制阈值取 0.3。

#### 3.2 实验对比及分析

针对炸点目标为了选择最优的金字塔特征融合方式，进行了不同模型的对比实验，结果如表 3 所示。AP<sub>50</sub> 表示当预测框和真值框交并比大于 0.5 时，认为是正样本，此时对应的炸点平均识别精度。AP 表示交并比取 [0.5, 0.65, …, 0.95] 得到的各个识别精

度的平均值。其中第 6、7 行等价于使用单一尺度特征的结构。实验结果表明：对于炸点目标识别，利用多尺度特征融合进行预测的结果好于单一尺度特征；其中，结合 P2~P6 尺度的特征信息能达到最好的预测效果，但由于 P6 层特征对于小目标作用不大，效果提升不明显，同时会带来运算量增加。实验 2—6 表明，增加低层信息能够有效提高炸点预测准确率；实验 1—4、6 表明，结合特征尺度数目增加，预测结果准确率越高，但当结合信息足够全面时，再增加其他尺度特征对炸点识别结果不会有太大影响。同时对比实验 6、7 表明，仅利用单层细节信息的模型不能达到很好的效果。

表 3 特征融合对比实验结果

序号	P2	P3	P4	P5	P6	AP <sub>50</sub>	AP <sub>75</sub>	AP
1	√	√	√	√	√	0.962	0.933	0.912
2	√	√	√	√		0.960	0.922	0.907
3		√	√	√		0.952	0.916	0.896
4		√	√			0.941	0.912	0.884
5		√		√		0.925	0.893	0.872
6		√				0.853	0.854	0.859
7			√			0.844	0.821	0.817

为了验证模型性能，选择两阶段的 Mask-RCNN 和单阶段的 YOLACT 2 种算法，利用相同的训练集和训练策略进行模型训练，在测试集上进行结果测试，以作为对比实验。Mask-RCNN 采用 ResNet101 作为基础网络，YOLACT 分别使用 Darknet53、ResNet101 作为基础网络，本算法采用 ResNet50-FPN 作为基础网络，速度、模型参数量及精度如表 4 所示。从参数量来看，本算法参数量仅为 29.42 M；从单张图像时间开销来看，是两阶段算法 Mask-RCNN 的 16.7 倍，同时快于其他对比算法，表明提出的算法在速度上更具优势，能够达到实时性的要求。

表 4 不同模型对比实验结果

模型	速度/fps	参数量/M	浮点运算量/G	AP <sub>50</sub>
Mask-RCNN	1.2	243.50	103.63	0.969
YOLACT-Darknet53	20.4	47.22	72.03	0.956
YOLACT-ResNet101	16.6	187.20	89.03	0.963
文中算法	21.2	29.42	54.61	0.952

图 6 展示了使用 P3~P5 尺度融合的算法炸点识别效果，通过质心计算标记出炸点中心位置。可以看到：对于小尺度的炸点，本模型能够较好地获取炸点区域并根据轮廓信息得到炸点位置。在实验环境下，本模型在炸点测试集上达到 95.2% 的 AP<sub>50</sub>，速度达到 21.2 fps，能够准确快速地得到炸点位置，满足炮兵对抗系统对准确率和实时性的要求。

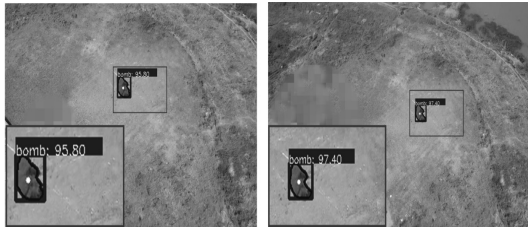


图 6 部分炸点捕捉示意

### 4 结束语

针对炮兵对抗训练系统中炸点捕捉难的问题，笔者提出一种基于 YOLACT 的快速炸点区域分割方法。针对炸点目标特性对特征提取网络特征融合方式及候选框参数进行修改和微调，提高了识别精度和速度；通过掩膜生成网络和预测分支对多个掩膜进行融合，实现了炸点目标识别及分割。实验结果表明：在构建的炸点数据集上，文中算法实现了快速准确地识别炸点目标，能够较好地解决炮兵对抗训练中的问题。

### 参考文献：

[1] 靳莹, 杨润泽. 声测定位技术的现状研究[J]. 电声技术, 2007, 31(2): 4-8.

[2] 羽佳. 主动式高速目标探测方法的研究[D]. 天津: 天津大学, 2004: 9-12.

[3] 赵艳丽. 弹道导弹雷达跟踪与识别研究[D]. 长沙: 国防科技大学, 2007: 5-7.

\*\*\*\*\*

(上接第 15 页)

[5] WANG R, LU B, HOU Y, et al. Passivity-based control for rocket launcher position servo system based on improved active disturbance rejection technology[J]. Advances in Mechanical Engineering, 2018, 10(3): 168781401876674.

[6] 魏立新, 赵默林, 范锐, 等. 基于改进鲨鱼优化算法的自抗扰控制参数整定[J]. 控制与决策, 2019, 34(4): 816-820.

[7] 胡达, 童仲志, 侯远龙, 等. 基于神经网络的 PMSM 分数阶积分滑模控制[J]. 兵工自动化, 2019(5): 81-85.

[8] 周维华. RBF 神经网络隐层结构与参数优化研究[D]. 上海: 华东理工大学, 2014: 23-28.

[9] 姜雪莹, 陶文华, 施惠元, 等. 基于 RBF 神经网络的非线性迭代预测控制[J]. 控制工程, 2019, 26(3): 454-460.

[10] 李匡成. 基于 RBF 神经网络的伺服系统自适应自抗扰控制[J]. 电气自动化, 2010, 32(2): 23-25.

[11] 袁川. 某型多管火箭炮伺服系统负载扰动抑制研究与

防科学技术大学, 2007: 5-7.

[4] 刘峰, 赵广伟, 王向军. 增强区域全卷积网络下的炸点检测方法研究[J]. 计算机辅助设计与图形学学报, 2019, 31(3): 412-420.

[5] 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, 2015, 39(6): 1137-1149.

[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]. European Conference on Computer Vision, 2016: 21-37.

[7] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection[EB/OL]. [2019-04-02]. <https://arxiv.org/abs/1904.01355>.

[8] KAIMING H, GEORGIA G, PIOTR D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 99: 1.

[9] BOLYA D, ZHOU C, XIAO F, et al. YOLACT: real-time instance segmentation[EB/OL]. [2019-04-04]. <https://arxiv.org/abs/1904.02689>.

[10] 王涵, 朱虹旭, 周双喜. 基于虚拟现实技术的对抗式指挥训练[J]. 国防科技, 2018, 39(1): 117-119.

[11] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Honolulu: The 31th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, July 22-25, 2017.

[12] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

\*\*\*\*\*

实现[D]. 南京: 南京理工大学, 2014: 46-51.

[12] 丁网芳, 刘星桥. 采用自抗扰控制的三电机同步协调系统[J]. 信息技术, 2018, 27(2): 22-27.

[13] HAGAN M T, MENHAJ M B. Training feedforward networks with the Marquardt algorithm[J]. IEEE Transactions on Neural Networks, 1994, 5(6): 989-993.

[14] WILAMOWSKI B M, CHEN Y, MALINOWSKI A. Efficient algorithm for training neural networks with one hidden layer[C]//International Joint Conference on Neural Networks. IEEE Xplore, 1999: 827-834.

[15] WILAMOWSKI B M, YU H. Improved Computation for Levenberg-Marquardt Training[J]. IEEE Transactions on Neural Networks, 2010, 21(6): 930-937.

[16] YU H, REINER P D, XIE T, et al. An Incremental Design of Radial Basis Function Networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25(10): 1793-1803.

[17] 韩京清. 自抗扰控制技术——估计补偿不确定因素的控制技术[M]. 北京: 国防工业出版社, 2008: 276-292.