

基于显著性目标检测的改进火焰检测算法

路茗^{1,2}, 谭锦钢^{1,2}, 张质懿¹, 陈明³, 何为^{1*}

¹中国科学院上海微系统与信息技术研究所无线传感网与通信重点实验室, 上海 201800;

²中国科学院大学, 北京 100864;

³中科院无锡高新微纳传感网工程技术研发中心, 江苏 无锡 214135

摘要 针对靶场炮弹定位方法定位精度差、实时性差且存在安全隐患等问题, 提出了一种基于显著性目标检测的炮弹火焰检测算法。首先针对数据集缺失问题, 构建了一个炮弹火焰数据集, 用于网络模型训练和推理。其次, 采用并列交叉的双分支 ResNet 为特征提取模块, 分别学习前景和背景语义信息, 并在该模块中引入空洞卷积和注意力机制网络, 提高感受野的同时使网络学习聚焦有用通道和空间位置的能力; 最后, 引入双向金字塔结构 (Bi-FPN) 融合浅层纹理信息及深层语义信息, 实现多尺度多阶段预测。实验结果表明, 所提算法在检测炮弹火焰的位置精确性、区域完整性和抗干扰性方面都明显优于现有算法, 可满足靶场日常掷弹定位训练需求。

关键词 图像处理; 炮弹定位; 显著性目标检测; 多尺度检测; 火焰检测

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202259.0410012

Improved Flame Detection Algorithm Based on Salient Target Detection

Lu Ming^{1,2}, Tan Jingang^{1,2}, Zhang Zhiyi¹, Chen Ming³, He Wei^{1*}

¹Key Laboratory of Wireless Sensor Networks and Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 201800, China;

²University of Chinese Academy of Sciences, Beijing 100864, China;

³Wuxi Hi-Tech Nano SensingNet R&D Center of Chinese Academy of Sciences, Wuxi, Jiangsu 214135, China

Abstract Aiming at the problems of poor positioning accuracy, bad real-time performance, and potential safety hazards of shell positioning method in the range, a flame detection algorithm based on salient object detection is proposed. First, in view of missing datasets, a shell flame dataset is constructed for network model training and reasoning. Second, the parallel and crossed two-branch ResNet is used as the feature extraction module to learn foreground and background semantic information respectively. Furthermore, dilated convolution and attention mechanism are introduced to improve the receptive field and synchronously enables the network to learn the ability of focusing on useful channels and spatial locations. Finally, the Bi-directional feature pyramid network (Bi-FPN) is introduced to fuse shallow texture information and deep semantic information to realize multi-scale and multi-stage prediction. Experimental results show that the proposed algorithm significantly outperforms the existing algorithms in terms of the accuracy, the regional integrity, and anti-interference, which is able to meet the needs of daily projectile positioning training in the shooting range.

Key words image processing; shell positioning; salient target detection; multi-scale detection; flame detection

收稿日期: 2021-03-03; 修回日期: 2021-03-29; 录用日期: 2021-04-02

基金项目: 国家重点研发计划 (2018YFC1505204)

通信作者: *wei.he@mail.sim.ac.cn

1 引言

随着武器装备不断向智能化方向发展,传感技术^[1]、通信技术、人工智能技术开始广泛应用于军事领域。在实际战场对抗中,炮弹落点位置是否准确是评估武器性能和操作者控制技术的重要参考指标之一。目前,国内靶场炮弹定位的主流方法还是人工测量^[2],即借助卷尺对炮弹落点与靶标位置进行实地测量。该方法存在效率低下、实时性差和未知安全隐患等问题。随着人工智能时代的到来,深度学习广泛应用于数字图像处理领域,包括目标检测、目标识别与追踪等,这也为靶场炮弹定位技术新一轮的革新带来希望。

炮弹落地爆破后主要呈现三种状态:烟尘、火焰及二者混合。在实战炮弹爆破目标位置检测中,因爆破产生的烟尘受风力影响会移动、扩散,而炮弹爆破产生的火焰几乎不受风力影响,更能准确表示炮弹落点位置,因此为了获得准确的炮弹爆破位置,炮弹火焰的检测显得尤为重要。早期火焰检测着重于研究火焰颜色、纹理等浅层信息,但是手工设计的特征描述对环境适应力差、鲁棒性弱,比如文献[3]研究火焰在RGB颜色空间和HIS颜色空间的色度及火焰动态特征来完成火焰检测;文献[4]利用火焰在YUV颜色空间上的色度和火焰运动特征来提取火焰像素;文献[5]利用YCbCr空间的特殊规则来区分火焰和非火焰的色度和亮度。随着卷积神经网络(CNN)的快速发展,基于深度学习的火焰检测方法也逐渐出现。起初,基于深度学习的火焰检测主要是利用火焰数据集对已有网络架构进行微调,使网络可检测复杂环境下的火焰。比如文献[6]基于GoogleNet进行微调、文献[7]对AlexNet架构进行微调。随后,部分学者对已有网络进行修改,利用火焰特性让网络架构更加贴合火焰检测任务。比如文献[8]对轻量级网络架构MobileNet进行修改,使其可以应用于工业环境中;文献[9]用优化过的AlexNet先在颜色空间上进行颜色分割,然后再进行特征提取和分类;文献[10]通过调整RetinaNet的训练策略和参数,使其在复杂环境下也能保持较高检测率;文献[11]使用K-means优化锚框大小后,再用YOLOv4检测火焰,可降低误检率;文献[12]将CNN和最大相关最小冗余结合,提高了火焰检测性能和泛化能力;文献[13]将全卷积网络与定位信度结合,提高了候选框定位精度及检测精度;文献

[14]利用双颜色空间特征提取前景特征后,用灰度共生矩阵获取纹理特征并提取在YCbCr空间的颜色特征,融合二者来提高火焰识别精度。综上所述,已有工作虽然都具备火焰目标实时检测能力,但数据集大多是由矩形框进行标注的,特别是火焰受外界因素影响变得极度不规则时,该矩形框内的背景信息容易影响网络检测精度,进而导致炮弹定位系统精准度和抗干扰能力不足。由此可见,表征炮弹落地位置的火焰检测算法性能仍需进一步提升。

显著性目标是人们观察图像时,容易捕捉到的图像中最为显著的部分。针对炮弹落地位置检测任务,因为有炮弹落点位置的先验,所以火焰区域总是人眼最为关注的区域,即可将其定义为显著性目标检测任务。传统显著性目标检测方法通常利用颜色、边缘、纹理等浅层特征的差异获得显著图,该方法泛化能力弱、针对性较强,不适用于复杂靶场环境中。近年来,基于深度学习的显著性目标检测得到迅速发展。文献[15]用全卷积网络R-FCN提取显著性目标的深层语义信息,获得显著图。文献[16]先用卷积网络获得粗略显著图,然后引入金字塔结构获得精准的显著图。文献[17]对前景和背景交替进行显著性检测和轮廓检测,从而获得显著图。文献[18]提出一种双向信息传递机制,该机制可增强语义信息及空间细节,获得目标边界清晰的显著图。文献[19]提出PICANets结构,该结构可以增强网络学习全局对比度的能力,提高显著检测性能。文献[20]提出BASNet架构和新损失函数,该结构可细化显著性目标边缘。文献[21]建立了双路耦合的特征融合网络,该网络可提高显著性目标检测性能。虽然部分显著性网络能获得较好的火焰分割结果,但在检测精度和速度上仍有可提升的空间。

因此,本文提出了一种基于显著性目标检测的新型火焰检测算法。首先,针对炮弹火焰定位任务中数据集缺失的问题,提出了一个炮弹火焰检测数据集。其次,以并列的两分支ResNet为特征提取有效途径,并在两支路中引入短连接支路共享部分语义信息,增强了网络的环境适应力,提高了网络鲁棒性。然后,在ResNet模块中引入空洞卷积和注意力机制网络,提高感受野的同时使网络学习聚焦有用通道和空间位置的能力,强化图像中火焰显著区域有效分割能力。最后,引入双向特征金字塔结构(Bi-FPN)融合浅层纹理信息及深层语义信息,提升多尺度特征信息在网络解码中的利用率。

2 基于显著性目标检测的火焰分割算法

2.1 所提算法流程

所提火焰检测算法主要基于显著性目标检测模型,包括模型训练(model training)和模型推理(model inference)。在模型训练阶段,首先构造火焰分割训练数据集,然后建立基于显著性目标检测算法的火焰检测模型,最后基于所构数据集输入状况,提炼模型输出的最优参数。在模型推理阶段,基于真实应用场景所获实时靶场图像构成的测试集进行显著性火焰检测并精准输出结果。所提算法流程如图 1 所示。

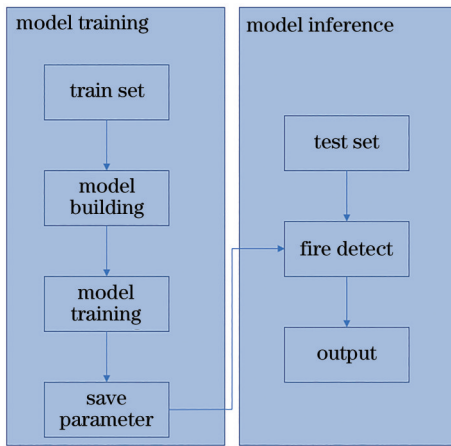


图 1 所提算法流程图

Fig. 1 Flow chart of proposed algorithm

2.2 改进的显著性火焰分割算法

2.2.1 改进的火焰分割模型

随着人类视觉系统和深度学习逐渐受到研究者的关注,基于神经网络的显著性目标检测逐渐成为主流方法,但在实时性和检测精度上仍有可提升的空间。为快速准确地捕捉到图像中火焰区域信息,提出了基于深度学习的显著性火焰分割模型,该模型主要包括特征提取网络(backbone)、短连接支路网络(short connection)及双向金字塔结构(Bi-FPN)3个模块,网络架构如图 2 所示。

从图 2 中可以看出,特征提取网络上下支路是对称的,且通过短连接支路共享语义信息,模拟人类视觉处理系统中数据交流方式,在输出结果前引入双向金字塔结构,将浅层纹理信息与深层语义信息融合,实现多尺度多阶段目标识别。模型输出结果为两张二值化显著图像,上路输出火焰显著图,白色区域为火焰区域,黑色区域为背景,下路输出

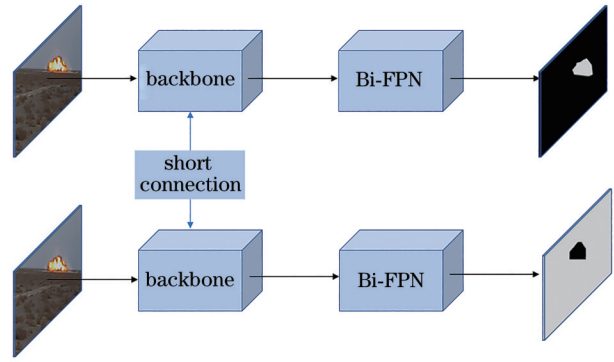


图 2 改进的火焰检测模型

Fig. 2 Improved fire detection model

背景显著图,白色区域为背景,黑色区域为火焰区域。

2.2.2 特征提取网络

传统显著性目标检测常常需要对比前景与背景的颜色、纹理等底层特征的差异,且野外靶场背景复杂多变,因此背景学习对于准确识别目标具有重要意义^[17]。常用特征提取模块基本是基于单列 CNN 架构的,比如 ResNet^[22]、VGG 等^[23]。在模型加深时,ResNet 中 residual block 结构可很好地解决退化问题,因此所提算法利用并行的 ResNet 架构交叉连接来模拟人类双眼,分别学习前景和背景信息^[18]。特征提取网络架构如图 3 所示。

由图 3 可知,特征提取网络的主体是两并列交叉的 ResNet 网络。ResNet 网络由多层 residual block 构成,可提取到多层多尺度特征,低层特征被卷积次数少,还保存着良好的空间位置信息,对定位显著性目标边界具有重要意义。而经过高度卷积后得到的高层特征具有丰富的语义信息,可定位显著性目标,增强网络泛化能力。人类视觉系统拥有极强的信息交流和数据处理能力,所提算法引入短连接支路模拟信息交流功能,达到增强边界信息和语义信息的目的,然后将增强后的特征图 C_2, C_3, C_4, C_5 输出到双向金字塔网络中,进而提升火焰检测的多尺度预测能力。

2.2.3 短连接支路

由文献[18]可知,边界信息及深层语义信息的增强可以提高显著性目标检测性能。浅层特征图中存在丰富空间位置及纹理信息,深层特征图中具有高级语义信息,为同时增强二者信息,提出一种新型短连接支路,架构如图 4 所示。

从图 4 中可以看出,在短连接支路中,先对上路特征进行空洞卷积^[24]扩大特征感受野,保存火焰边

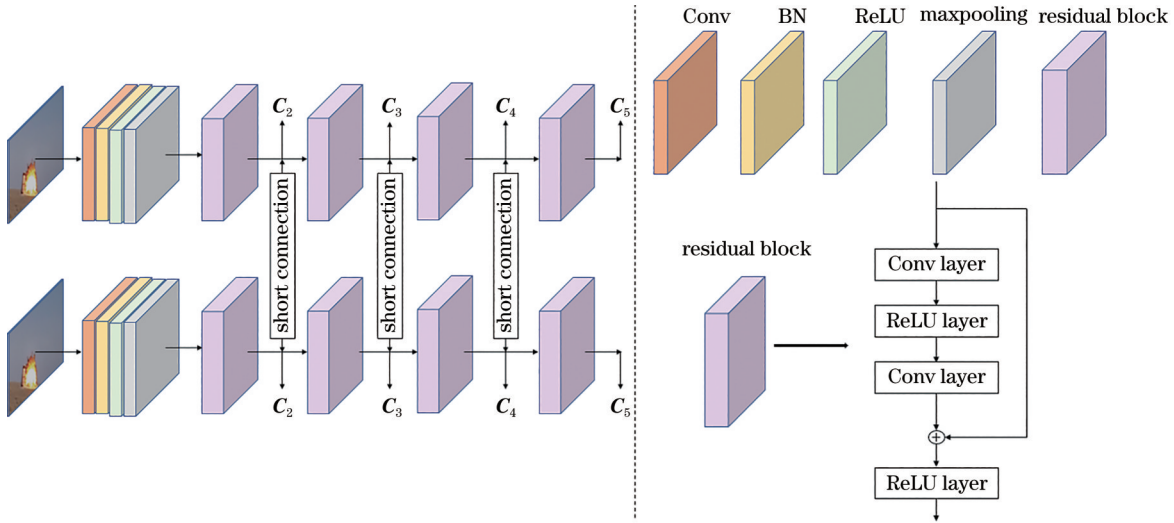


图 3 特征提取网络

Fig. 3 Feature extraction network

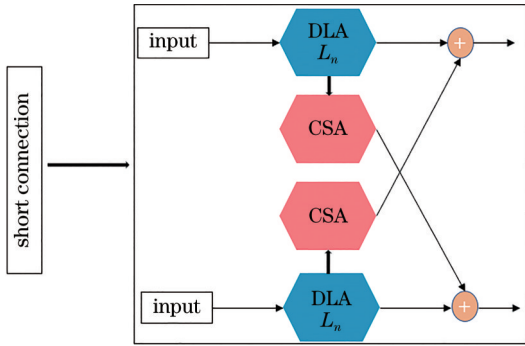


图 4 短连接支路网络

Fig. 4 Short connection branch network

界信息。然后引入注意力机制(CSA)^[25],使双路 ResNet 聚焦于目标信息丰富的通道和空间位置。与普通卷积相比,空洞卷积增加了空洞率指标,即在卷积操作中增加零填充,可在同分辨率特征

图下,提高卷积感受野,进一步解决信息丢失问题。真实卷积核尺寸与等效卷积核尺寸之间的关系为

$$K = (d - 1)(k - 1) + k, \quad (1)$$

式中: d 为空洞率; k 为真实空洞卷积核尺寸; K 为等效卷积核尺寸。从(1)式可以看出,空洞卷积的等效卷积核尺寸 K 更大,表示感受野范围更大。考虑到不同尺度特征图大小不一,感受野扩大范围有限,提出一种新的空洞卷积分布机制(DLA),网络架构如图 5 所示。

从图 5 中可以看出,DLA 中空洞卷积层的数量 $L_n, n=1, 2, 3$ 与 ResNet 中 block 层有关。浅层 block 输出特征图较大,则需要多次空洞卷积操作才可扩大特征感受野;深层 block 输出特征图较小,只需要进行一次空洞卷积操作即可。

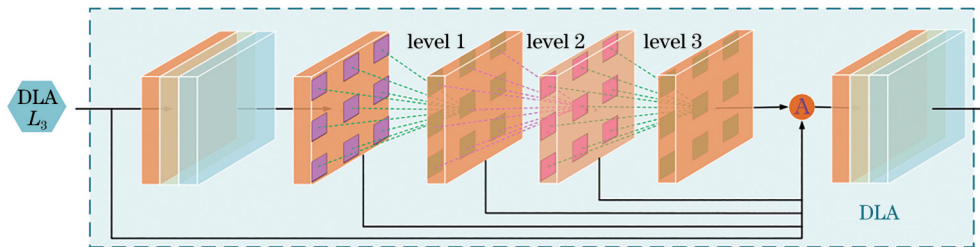


图 5 空洞卷积模块

Fig. 5 Dilated convolution module

注意力机制在神经网络中广受关注,文献[26]提出的通道注意力机制更关注特征通道之间的关系,即对特征通道进行再标定操作,使网络聚焦于更“有用”的通道信息,可描述为

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (2)$$

$$S_c = \sigma[F_2 \delta(F_1 Z_c)], \quad (3)$$

$$X = S_c Z_c, \quad (4)$$

式中： H, W 分别为图像的长宽； u_c 为输入图像 u 第 c 通道特征图； Z_c 为对特征图 u 进行全局平均池化操作得到的压缩后的特征向量； δ 表示 ReLU 函数； F_1, F_2 表示两层全连接层； σ 表示 Sigmoid 函数； S_c 代表第 c 个通道的权重； X 为得到的更关注“有用”通道的特征图。特征向量 Z_c 经过全连接层和 ReLU 层后，降低参数数量的同时使模型具有更优的非线性。进而，该特征向量经过全连接层和 Sigmoid 层，进行升维操作并将该特征向量归一化在 $[0, 1]$ 范围内，从而得到代表第 c 个通道的权重 S_c 。最后，将该权重信息加权到原图 Z_c 上，即可得到更关注有用通道的 X 特征图。

空间注意力机制^[22]最早提出于图像处理任务中，主要模拟人类大脑系统的数据处理机制，将注意力更多关注于目标相关的区域内，可描述为

$$Y = \sigma[F(X)]X, \quad (5)$$

式中： F 代表 1×1 卷积操作。首先，将特征图 X 压

缩成 $W \times H \times 1$ 维，再经过 Sigmoid 函数进行归一化，每一个空间位置都代表该位置的一个权重信息，归一化后的权重特征图与原图 X 进行加权，得到特征图 Y 即完成对空间位置的标定工作。

通道注意力和空间注意力二者是相辅相成的关系，为了使网络同时学习到特征通道和空间位置的信息，所提改进注意力网络 CSA 的具体结构如图 6 所示。

从图 6 中可以看出，改进的注意力网络是将通道注意力和空间注意力二者串行连接所得的，输入特征图 I_1 先通过通道注意力(channel attention)模块完成特征通道的选择，然后通过空间注意力(spatial attention)模块实现空间位置的筛选工作，所得特征图 I_3 与 I_1 尺度一致，但比 I_1 更有意义。综上所述，通道注意力和空间注意力使网络自主学习每个通道和每个空间位置的权重信息，权重越高表示该通道或空间位置对目标检测任务越重要，让网络拥有自主学习到聚焦有用通道和空间位置的能力。

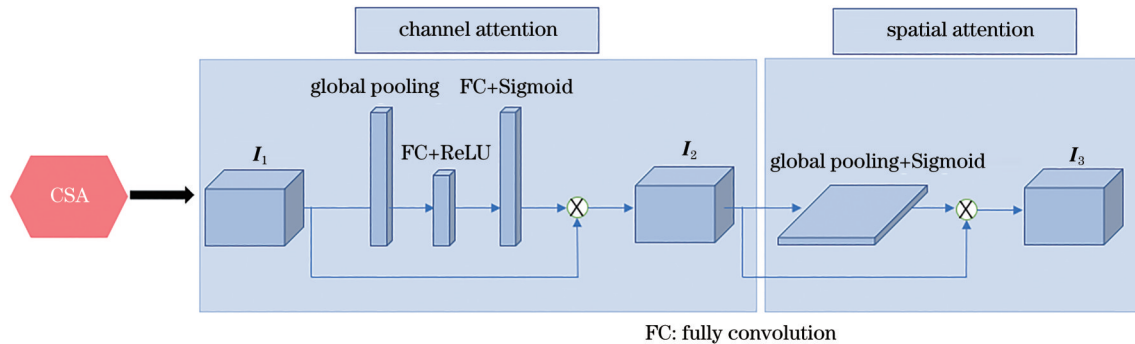


图 6 注意力机制网络

Fig. 6 Attention mechanism network

2.2.4 双向金字塔结构

传统的 FPN^[27] 结构只包含相邻两层的语义信息，较深层的语义信息不包含颜色、纹理和空间位置等目标浅层特征，容易造成定位准确性差的问题。为了弥补 FPN 结构的不足，所提算法增加 1 个自下到上分支(D2T)，并将该分支与自上到下分支(T2D)融合，输出 4 张和标签同样大小的、归一化到 $[0, 1]$ 的预测特征图。所提 Bi-FPN 包含 T2D 和 D2T 两个部分，如图 7 所示，其中 transpose convolution network (TCN) 用反卷积实现特征图上采样，channel downsample (CD) 的作用是将输入特征图参数数量减半。

从图 7 中可以看出，双向特征金字塔输入为 C_2, C_3, C_4, C_5 特征图，T2D 是卷积神经网络的前馈

计算，将高层的具有丰富语义的特征图进行 2 倍上采样后得到 T_5, T_4, T_3 。为增强对有用通道及位置的学习，对 T_5, T_4, T_3 再进行一次注意力机制操作后，与相邻特征图进行融合。T2D 的网络架构与 FPN 架构中自上向下网络一致，但该 FPN 架构是从深到浅累积信息的，导致输出只包含本身层及更深层的特征，因此很难获得更优的检测结果。为获取更优的检测结果，所提算法新增 D2T 分支，从浅层包含丰富的纹理及空间位置的特征图 C_2 开始，逐步经过注意力模块后与更深层的特征拼接，从而累计从浅到深的特征信息。所提 Bi-FPN 结构使得每一个输出阶段都能利用到完整的深层语义信息和浅层纹理位置信息，因此在每一阶段都可以得到更精确的分割结构。最后对预测得到的 A_2, A_3, A_4, A_5 相

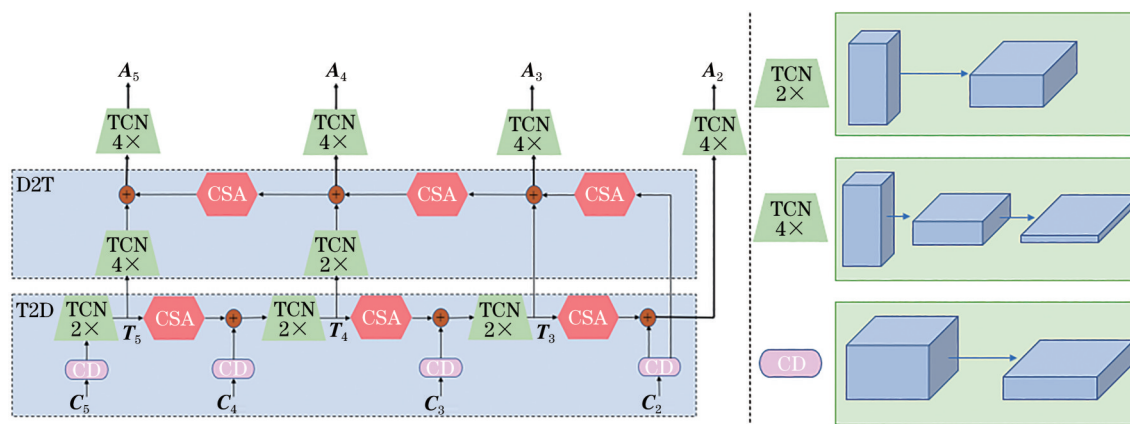


图 7 双向金字塔网络
Fig. 7 Bi-FPN

加后平均,获得最后的前景预测结果和背景预测结果。

3 网络训练

3.1 构建数据集

炮弹火焰检测的目的是获取准确的炸点位置,基于这一任务需求,炮弹火焰数据集需要满足两个条件:1)有足够多的炮弹火焰数据;2)准确标注炮弹火焰区域。目前存在的火焰有效数据集有两大缺点:1)这些数据集是通过谷歌、百度等搜索引擎从各地下载的,与炮弹火焰状态差异大,例如森林火灾、建筑火灾、蜡烛等不同来源的火焰数据;2)人工标注

仅可满足锚框类目标检测的标定要求,且标注效果差,容易造成训练的网络检测精度不佳等情况。为了解决上述问题,所提算法基于靶场真实炮弹爆炸视频获取炮弹火焰数据集,而且针对锚框类和分割类两种训练目的,用不同标注方法对相同数据集进行标注,用锚框标注时尽量准确、完整地包括火焰,显著性目标检测标注方法则尽量勾勒出火焰形状,两个数据集分别用于目标检测和显著性目标检测。训练集共 1339 张图片,包括 1159 张炮弹爆炸时火焰图片、180 张日出、日落及灯光的图片;测试集共 400 张图片,包括 300 张炮弹火焰照片和 100 张红色和火色的图片。部分数据集如图 8 所示。

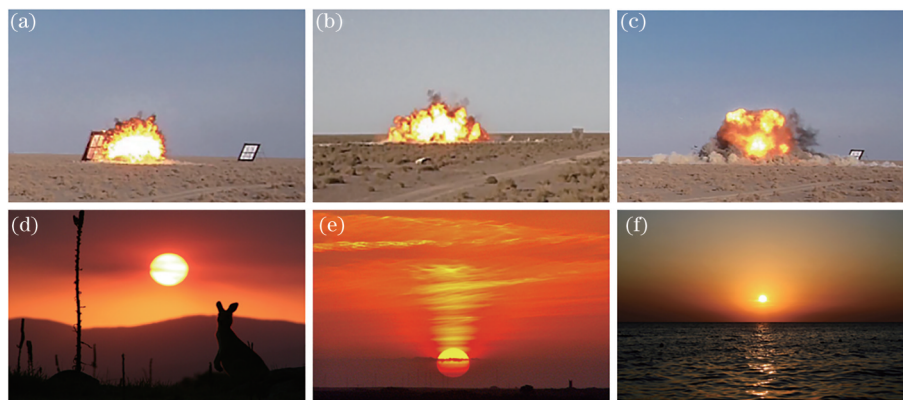


图 8 部分样本数据集。(a)~(c)正样本;(d)~(f)负样本
Fig. 8 Part of sample dataset. (a)~(c) Positive samples; (d)~(f) negative samples

图 8 中,第 1 行是部分正样本,炮弹火焰具有密集性、紧凑性的特征;第 2 行是类似火焰的落日、灯光等负样本照片。加入负样本进行训练可提高网络分类准确性和鲁棒性。深度神经网络模型性能与数据集中数据量的大小有极大的关系,内容丰富的数据集训练所得模型检测性能更优,小样本量数

据集即使在训练集上获得优异性能,也容易因为过拟合造成在测试集上性能欠优。扩增数据集是解决过拟合问题的重要途径,除了每次测试后收集更多炮弹火焰图像以外,还利用图像随机翻转、随机旋转、亮度和对比度调整、随机缩放等数据增强方式扩增数据集。

3.2 评价指标

查准率 (precision) 和查全率 (recall) 是评估网络性能的重要指标, 表达式分别为

$$R_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (6)$$

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (7)$$

式中: N_{TP} 为预测正确率; N_{FP} 为误检率; N_{TN} 为背景正确率; N_{FN} 为漏检率。查准率表示在所有预测为正的样本中预测正确的比例; 查全率表示在所有正样本中预测正确的比例。尽管网络性能评估中常用到查准率和查全率, 但二者均不能较为全面地对显著性目标检测网络进行评估。因此, 所提算法采用的评估指标是显著性目标检测中常用的性能指标 F-measure 和平均绝对误差 (MAE), 表达式为

$$F_{\beta} = \frac{(1 + \beta^2)R_{\text{precision}} \times R_{\text{recall}}}{\beta^2 R_{\text{precision}} + R_{\text{recall}}}, \quad (8)$$

$$E_{\text{MAE}} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W |\mathbf{P}(i, j) - \mathbf{G}(i, j)|, \quad (9)$$

式中: β 表示权重。 \mathbf{P} 和 \mathbf{G} 分别代表归一化后的预测图及标签图; $\mathbf{P}(i, j)$ 和 $\mathbf{G}(i, j)$ 代表预测图 \mathbf{P} 和标签图 \mathbf{G} 中在位置 (i, j) 的像素值, 它们的范围都在 $[0, 1]$; H 、 W 分别为图的长宽。权重 β 是为了均衡查全率和查准率两个指标, 实验中取 $\beta^2 = 0.3$, 表示增加了查准率的权重, 表示网络训练过程中认为查准率更为重要。 F_{β} 值越高, 表示网络检测能力越强, 反之越弱。 MAE 主要评估两张图之间的差异, MAE 值越小, 说明两张图越相似。

3.3 实验结果与分析

为探究不同深度 ResNet 网络结构对检测性能的影响, 分别对 ResNet18、ResNet34、ResNet50 进行训练和测试, 评价指标为 MAE 和 F-measure。由于测试集中加入部分负样本照片, 为了获得较为准确的 F-measure, 将数据集中有前景和没有前景的照片分开进行评价。对于有前景的照片, 直接计算 F-measure, 记为 F_{β_fg} ; 对于没有前景的图片, 评估背景准确度记为 F_{β_bg} 。实验结果如表 1 所示

表 1 不同深度特征提取网络的性能分析

Table 1 Performance analysis of feature extraction networks with different depth

Parameter	ResNet-18	ResNet-34	ResNet-50
MAE	0.0101	0.0138	0.0104
F_{β_fg}	0.8812	0.8815	0.8943
F_{β_bg}	0.9010	0.8802	0.8734

从表 1 中可以看出, 不同深度 ResNet 在训练相同轮次后, 平均误差均小于 0.015, 前景检测准确度、背景检测准确度均较高, 均能达到 0.87 以上。但 ResNet 的加深仅略提高检测性能, 对网络模型影响不大, ResNet-18 模型更轻量, 更适合应用于靶场射击实施定位系统中。

为了检验所提算法的有效性和效果准确性, 将所提算法和 YOLOv4^[11]、RetinaNet^[10]、BASNet^[16]、PICANet^[17] 的检测性能和推理速度进行对比。对比结果如表 2 和表 3 所示。

表 2 所提算法与 YOLOv4、RetinaNet 性能对比

Table 2 Performance comparison between proposed algorithm and YOLOv4, RetinaNet

Parameter	Proposed algorithm	YOLOv4	RetinaNet
Precision	0.890	0.900	0.802
Recall	0.880	0.880	0.773
F_{β}	0.884	0.889	0.788

表 3 所提算法与 BASNet、PICANet 性能对比

Table 3 Performance comparison between proposed algorithm and BASNet, PICANet

Parameter	Proposed algorithm	BASNet	PICANet
F_{β}	0.881	0.877	0.861
MAE	0.010	0.013	0.023

从表 2 中可以看出, 所提算法的查准率、查全率和 F_{β} 都远优于 RetinaNet 但略小于 YOLOv4, 但 YOLOv4 预测出来的是包含火焰的方框, 而所提算法则需要精确地将火焰分割出来, 因此挑战和难度都更大。另外, YOLOv4 预测出来的矩形框容易包括大量的背景信息, 难以获得精确的炸点位置, 相对而言, 所提算法可以精确定位火焰位置, 进而确定爆炸点。因此所提算法在实际炮弹爆炸点定位上的效果远远超过 YOLOv4, 更具有实际意义。

从表 3 中可以看出, 所提算法与常用显著性目标检测网络 BASNet、PICANet 相比, F_{β} 值和平均绝对误差都相对更优, 表明所提算法可以提高显著性目标检测性能。

为验证所提算法在靶场炮弹实时定位环境中的适用性, 与检测性能较优的 YOLOv4、BASNet 在检测速度和模型大小上进行了对比, 为准确评估不同模型的模型推断能力, 在相同的硬件设备 (GPU 为 Tesla k40c) 上进行训练和测试。实验结果如表 4 所示。

表 4 不同算法推理速度和模型大小对比

Table 4 Comparison of reasoning speed and model size of different algorithms

Parameter	Proposed algorithm	YOLOv4	RetinaNet	BASNet	PICANet
Detection speed / (frame·s ⁻¹)	4	4	4.5	2	3
Model size / MB	121.5	256	145.7	348.5	188.9

从表 4 中可以看出, YOLOv4 作为单阶段目标检测的代表, 在检测速度上得到一致好评, 所提算法与 YOLOv4 检测速度相当, 相比于 BASNet, 检测速度提高近 1 倍。在模型大小对比上, 所提算法大小为 121.5 MB, 仅为 YOLOv4 的一半, 略高于 BASNet 模型的三分之一。所提算法在检测精度和速度上取得了良好的平衡, 更适于靶场炮弹实时定位系统中。

测试集中包括类似火焰的目标, 为了体现所提算法在分类此类目标时的优势, 与 YOLOv4 进行了对比, 实验结果如表 5 所示,

从表 5 中可以看出, YOLOv4 的误检率和漏检率分别为 0.200 和 0.067, 说明在 100 张负样本中,

表 5 所提算法与 YOLOv4 检测结果对比

Table 5 Comparison of detection results between proposed algorithm and YOLOv4

Parameter	Proposed algorithm	YOLOv4
N_{FP}	0.020	0.200
N_{FN}	0.003	0.067
N_{TP}	0.980	0.800
N_{TN}	0.997	0.933

误检火焰的有 20 张, 在 300 张火焰图片中漏检 20 张。所提算法的误检率和漏检率分别为 0.020 和 0.003, 表明所提算法的抗干扰能力更强。

为了更直观地体现所提算法的检测结果, 将所提算法和 YOLOv4、BASNet 的检测结果进行了对比, 结果如图 9 所示。

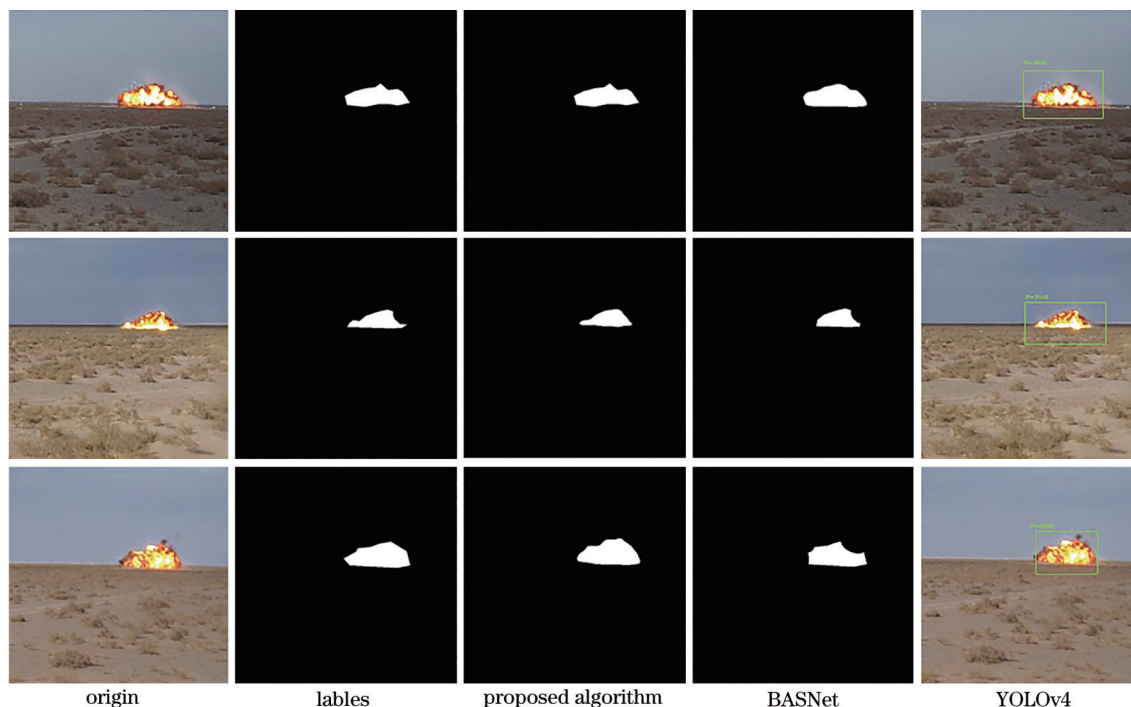


图 9 部分检测结果。

Fig. 9 Partial test results

在图 9 中, 第 1、2 列图像分别是火焰原图和标签, 第 3 列是所提算法的预测结果, 第 4 列是 BASNet 的分割结果, 第 5 列是 YOLOv4 的检测结果。所提算法火焰分割结果与 BASNet 的分割结果类似, 但可以获取更多火焰边缘信息, 获得的火焰

区域更为完整; YOLOv4 虽然可以圈出完整的火焰模型但是存在许多背景信息, 难以准确定位火焰位置。综上所述, 所提算法检测效果优于 YOLOv4、BASNet。

为验证所提算法中各个模块对检测性能的影响

响,在所提火焰数据集上进行了消融实验,实验结果如表 6 所示。

从表 6 可以看出,所提算法引入的各模块均在一定程度上提高了火焰检测性能。双向金字塔结构将浅层纹理信息与深层语义信息融合后,平均绝对误差降为原来的一半, F-measure 提高 0.0442。

引入空洞卷积扩大特征感受野后, F-measure 提高 0.0137。在特征提取模块及在金字塔中加入注意力机制后,平均绝对误差降低 0.022, F-measure 提高 0.0182。将双层 ResNet 网络架构变成单层 ResNet 后,平均绝对误差增加 0.022,而 F-measure 降低 0.034,表明双层网络架构的检测结果更优。

表 6 在火焰数据集上的消融实验
Table 6 Ablation experiments on flame dataset

Single ResNet	Double	Bi-FPN	Dilated Conv	Attention mechanism	MAE	F_{β}
	✓				0.085	0.8079
	✓	✓			0.043	0.8521
	✓	✓	✓		0.032	0.8658
	✓	✓	✓	✓	0.010	0.8840
✓		✓	✓	✓	0.032	0.8498

4 结 论

提出了一个新的火焰检测数据集并基于该数据集设计了一种有效、高效且鲁棒的火焰检测算法。首先,所提两分支 ResNet 架构能有效解决背景学习不足的问题,提高火焰分割精度。其次,利用注意力机制学习有用通道和位置,可以增强特征表达能力及网络抗干扰能力,进而提高检测精度。最后,所设计的双向特征金字塔解码结构和空洞卷积模块解决了图像分割边缘信息丢失的问题,提高了火焰分割完整性。综上所述,研究结果为军事掷弹训练及实际战场对抗提供有力技术支撑,在靶场对炮弹火焰快速定位中具有重要的工程实用价值。

参 考 文 献

- [1] Qu B N, Jiang P, Zhao L Y, et al. Passive direction finding method for shells based on short baseline sensor array[J/OL]. Journal of Aeronautics: 1-15 [2021-02-26]. <http://kns.cnki.net/kcms/detail/11.1929.V.20210225.1523.028.html>.
屈秉男, 蒋平, 赵鲁阳, 等. 基于短基线传感器阵列的炮弹被动测向方法[J/OL]. 航空学报:1-15 [2021-02-26]. <http://kns.cnki.net/kcms/detail/11.1929.V.20210225.1523.028.html>.
- [2] Qin X Y, Wang X F, Chen P, et al. Artillery blast point detection based on adaboost algorithm[J]. Acta Armamentarii, 2012, 33(6): 682-687.
秦晓燕, 王晓芳, 陈萍, 等. 基于 Adaboost 算法的炮弹炸点检测[J]. 兵工学报, 2012, 33(6): 682-687.
- [3] Chen T H, Wu P H, Chiou Y C. An early fire-detection method based on image processing[C]//

2004 International Conference on Image Processing, 2004. ICIP'04, October 24-27, 2004, Singapore. New York: IEEE Press, 2004: 1707-1710.

- [4] Marbach G, Loepfe M, Brupbacher T. An image processing technique for fire detection in video images [J]. Fire Safety Journal, 2006, 41(4): 285-289.
- [5] ÇelikTurgay, Demirel H. Fire detection in video sequences using a generic color model[J]. Fire Safety Journal, 2009, 44(2): 147-158.
- [6] Muhammad K, Ahmad J, Mehmood I, et al. Convolutional neural networks based fire detection in surveillance videos[J]. IEEE Access, 2018, 6: 18174-18183.
- [7] Muhammad K, Ahmad J, Lv Z, et al. Efficient deep CNN-based fire detection and localization in video surveillance applications[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019, 49(7): 1419-1434.
- [8] Muhammad K, Khan S, Elhoseny M, et al. Efficient fire detection for uncertain surveillance environment [J]. IEEE Transactions on Industrial Informatics, 2019, 15(5): 3113-3122.
- [9] Zhong Z, Wang M J, Shi Y K, et al. A convolutional neural network-based flame detection method in video sequence[J]. Signal, Image and Video Processing, 2018, 12(8): 1619-1627.
- [10] Jiang Y, Bai Y. Flame detection based on RetinaNet deep learning model[J]. Natural Science Journal of Hainan University, 2019, 37(4): 306-312.
江洋, 白勇. 基于 RetinaNet 深度学习模型的火焰检测研究[J]. 海南大学学报(自然科学版), 2019, 37(4): 306-312.

- [11] Yan X, Wu Z P, Lei B J. Research on video flame smoke detection based on deep learning[J]. *Information & Communications*, 2020, 33(9): 70-72.
颜洵, 吴正平, 雷帮军. 基于深度学习的视频火焰烟雾检测方法[J]. *信息通信*, 2020, 33(9): 70-72.
- [12] Li Z R, Wang H Q, Hu Y, et al. Flame image detection method based on deep learning with maximal relevance and minimal redundancy[J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 101015.
李梓瑞, 王慧琴, 胡燕, 等. 基于深度学习和最大相关最小冗余的火焰图像检测方法[J]. *激光与光电子学进展*, 2020, 57(10): 101015.
- [13] Zhang H, Yan Y Y, Liu Y A, et al. Fire detection method based on localization confidence and region-based fully convolutional network[J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 201021.
张鸿, 严云洋, 刘以安, 等. 基于定位置信度和区域全卷积网络的火焰检测方法[J]. *激光与光电子学进展*, 2020, 57(20): 201021.
- [14] Tan Y, Xie L B, Feng H W, et al. Flame detection algorithm based on image processing technology[J]. *Laser & Optoelectronics Progress*, 2019, 56(16): 161012.
谭勇, 谢林柏, 冯宏伟, 等. 基于图像的火焰检测算法[J]. *激光与光电子学进展*, 2019, 56(16): 161012.
- [15] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks [C]//*Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016*, December 5-10, 2016, Barcelona, Spain. [S.l.: s.n.], 2016: 379-387.
- [16] Wang T T, Borji A, Zhang L H, et al. A stagewise refinement model for detecting salient objects in images[C]//*2017 IEEE International Conference on Computer Vision (ICCV)*, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4039-4048.
- [17] Wu R M, Feng M Y, Guan W L, et al. A mutual learning method for salient object detection with intertwined multi-supervision[C]//*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 8142-8151.
- [18] Zhang L, Dai J, Lu H C, et al. A bi-directional message passing model for salient object detection [C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 1741-1750.
- [19] Liu N, Han J W, Yang M H. PiCANet: learning pixel-wise contextual attention for saliency detection [C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3089-3098.
- [20] Qin X B, Zhang Z C, Huang C Y, et al. BASNet: boundary-aware salient object detection[C]//*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 7479-7489.
- [21] Tan J G, Chen L L, Du L, et al. BVMSOD: bionic vision mechanism based salient object detection[C]//*2019 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, September 18-20, 2019, Munich, Germany. New York: IEEE Press, 2019: 335-339.
- [22] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [23] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2021-02-25]. <https://arxiv.org/abs/1409.1556>.
- [24] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 833-851.
- [25] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [26] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [27] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.