

基于融合混合域模块与空洞卷积的红外小目标检测

曲海成, 王馨心*, 欧阳俊

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105

摘要 针对红外图像中小目标分布稀疏、像素占比小, 现有红外小目标检测算法易受强噪声干扰, 对其精度和泛化性有较大影响的问题, 提出一种基于上下文信息融合与视觉显著性的红外小目标检测算法。首先, 主干网络采用编码-解码的方式构建, 其中, 编码层为空洞卷积所堆叠的全卷积神经网络, 对输入进行特征提取; 然后, 通过与解码层逐层跳跃拼接的方式, 实现跨层间的特征融合, 提取具有强语义和强位置的特征信息; 最后, 将提取特征输入混合域模块, 利用其中的通道注意力机制和空间注意力机制提高对小目标的特征权重, 以增强对背景的抑制。通过空洞卷积结合跨层融合以及利用混合域模块带来的视觉显著性, 所提算法在复杂背景下优于目前典型算法, 与当前最优算法相比, F_{measure} 平均提升了 10%, 运行效率加快 40%, 并且在检测率和虚警率指标上都有显著改善。

关键词 图像处理; 空洞卷积; 混合域; 注意力机制; 全卷积; 特征融合

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP213224

Infrared Small-Target Detection Based on Hybrid Domain Module and Hole Convolution

Qu Haicheng, Wang Xinxin*, Ouyang Jun

College of Software, Liaoning Technical University, Huludao 125105, Liaoning, China

Abstract To address the problems of sparse distribution of small targets in infrared images is sparse, the proportion of pixels is small, and existing infrared small-target detection algorithms are vulnerable to strong noise interference, which significantly impact their accuracy and generalization, an infrared small-target detection algorithm based on context information fusion and visual saliency is proposed. First, the backbone network is constructed by an encoding-decoding method, in which the encoding layer is a full convolutional neural network stacked by hole convolution, and the input features are extracted. Then, feature fusion between different layers is realized through layer-by-layer skip splicing with the decoding layer, and feature information with strong semantics and strong location is extracted. Finally, the extracted features are input into the mixed domain module, and the channel attention and spatial attention mechanisms are used to improve the feature weight of small targets to enhance the background suppression. Through hole convolution combined with cross-layer fusion and visual saliency provided by the hybrid domain module, the proposed algorithm is demonstrated to be superior to the current typical algorithm for complex backgrounds. Compared with the comprehensive optimal algorithm, F_{measure} is improved by 10% on average, operation efficiency is increased by 40%, and detection and false alarm rate indicators are improved significantly.

Key words image processing; cavity convolution; mixed domain; attention mechanism; full convolution; feature fusion

1 引言

红外探测是一种对物体进行定位、检测与识别的技术, 相较于雷达成像、可见光成像等具有隐蔽性、全时段和抗干扰能力强这 3 种优势。物体温度越高, 所对应的红外波段能量^[1]越大。根据这一特性, 红外探

测广泛应用于各个领域, 例如: 在民用领域包括气体泄漏检测、火灾预警、目标检测跟踪等; 在军用领域则包含对敌对目标的侦测、预警、跟踪等。

传统单帧红外弱小目标检测算法利用检测目标与背景存在灰度、结构等特征差异的特性设置检测算子^[2], 从而实现对目标的检测。比较典型的算法有

收稿日期: 2021-12-14; 修回日期: 2022-01-05; 录用日期: 2022-01-28; 网络首发日期: 2022-02-08

基金项目: 辽宁省教育厅科学研究项目(LJ2020QNL013)

通信作者: *wxxwang@foxmail.com

local contrast measure (LCM)^[3]、absolute directional mean difference (ADMD)^[4]、multiscale local homogeneity measure (MLHM)^[5]、multiscale patch-based contrast measure (MPCM)^[6] 以及 local intensity and measure (LIG)^[7] 等,这类算法的不足之处在于对背景强边缘和尖锐噪声的抑制效果较差。红外图像噪点较多,可通过设置图像背景抑制算法实现噪点过滤,从而检测目标^[8]。典型算法有最大中值滤波算法^[9]、top-hat 滤波算法^[10]。然而这类算法受限于人工设计特征的缺点,对噪声过滤能力有限且对复杂背景下小目标^[11-12]的检测效果较差。部分学者根据图像数据信息,通过低秩稀疏矩阵分解查找低维子空间结构,使用预先设定的矩阵分解算法来显示数据结构,实现小目标的检测工作,如 infrared patch-image model (IPIM)^[13] 和 low-rank and sparse matrix decomposition (LRSMD)^[14]。因为矩阵分解需要迭代进行,所以此类算法计算复杂、占用硬件资源较多且耗时较长。

传统单帧红外弱小目标检测算法高度依赖人工设计特征,算法设计复杂,并且在复杂背景下易受噪声^[15]和强边缘信息干扰。而深度学习由于具有样本量大、检测精度高等优点,近些年使用深度学习技术从事小目标检测研究成为了主流。学者普遍使用全卷积神经网络 (FCN)^[16] 及其变种,如引入池化索引的 SegNet^[17] 算法、使用跳跃拼接架构的 UNet^[18] 以及 UNet++^[19] 算法等,并将其应用于小目标检测。除此之外,也有学者将生成对抗网络应用于小目标检测工作,如 MDvsFA_cGAN^[20] 算法。这些深度学习算法的检测能力与背景抑制效果都要优于传统算法。

本文主要研究如何使用深度学习技术对单帧红

外图像进行小目标检测。由于探测距离过远,往往所要检测的目标在成像上尺寸占比非常小,在图像中呈现为一个点源^[21-22],并且受限于温度决定红外波段能量的特性,故而当目标所辐射的能量波段较小时,与背景的分度度较低。因此,红外探测技术是一项非常具有挑战性的工作。当前深度学习技术如生成对抗网络^[23]系列,往往伴随一个或多个生成器和判别器,网络结构设计复杂,因此,网络显存占用较大且运行耗时较长。而以 FCN^[24] 结合滤波为代表的红外小目标检测算法,受限于网络结构设计,感受野较小,对特征的提取及复用能力较弱,并且滤波操作对复杂背景的红外图像过滤噪声能力较差。根据当前存在的难点,本文提出结合空洞卷积逐层跨越连接的“编码-解码”设计方式,利用空洞卷积提高网络层感受野,逐层跨越连接进行各层之间的特征融合,增强语义信息和空间位置信息。之后使用混合域,在通道和空间上利用注意力机制放大重要特征信息,并且抑制尖锐噪声和强边缘。结合这些模块的使用,网络的层数和通道数无须设置过于复杂就能较好实现对小目标的特征信息提取。通过对比实验可知,所提端到端深度网络在运行效率以及综合评价指标上均比其他算法有提升。

2 所提算法

2.1 检测方案

所提算法检测方案如图 1 所示。首先,利用特征提取模块 infrared small target detection network (ISTDN) 对输入图 O 进行背景抑制和目标增强,输出特征图 S_1 。

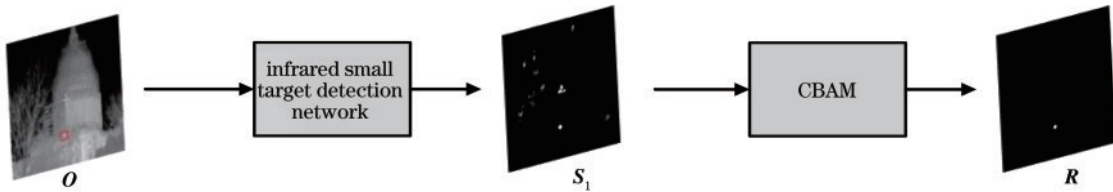


图 1 所提算法检测方案

Fig. 1 Detection scheme of proposed algorithm

然后,将特征图 S_1 输入混合域模块 convolutional block attention module (CBAM),通过其包含的通道注意力机制和空间注意力机制,对其进行视觉显著性分析:将显著特征权重放大,同时抑制尖锐噪声和强边缘特征,提高目标特征占比。最后,利用主干网络和混合域模块对图像信息的特征进行提取及过滤,得到准确率较高的结果 R 。计算公式如下:

$$S_1 = w \times O + b_{\text{bias}}, \quad (1)$$

$$R = \text{CBAM}(S_1), \quad (2)$$

式中: w 是通过特征提取模块生成的权重; b_{bias} 是偏置项。

图 2 是典型 FCN 检测方案。在编码器中将特征提取网络分成若干阶段,在解码器中将不同分支的特征信息融合并进行反卷积,输出结果 R 。图 3 网络是

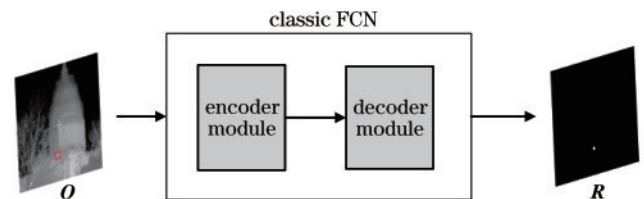


图 2 典型 FCN 检测方案

Fig. 2 Detection scheme of typical FCN

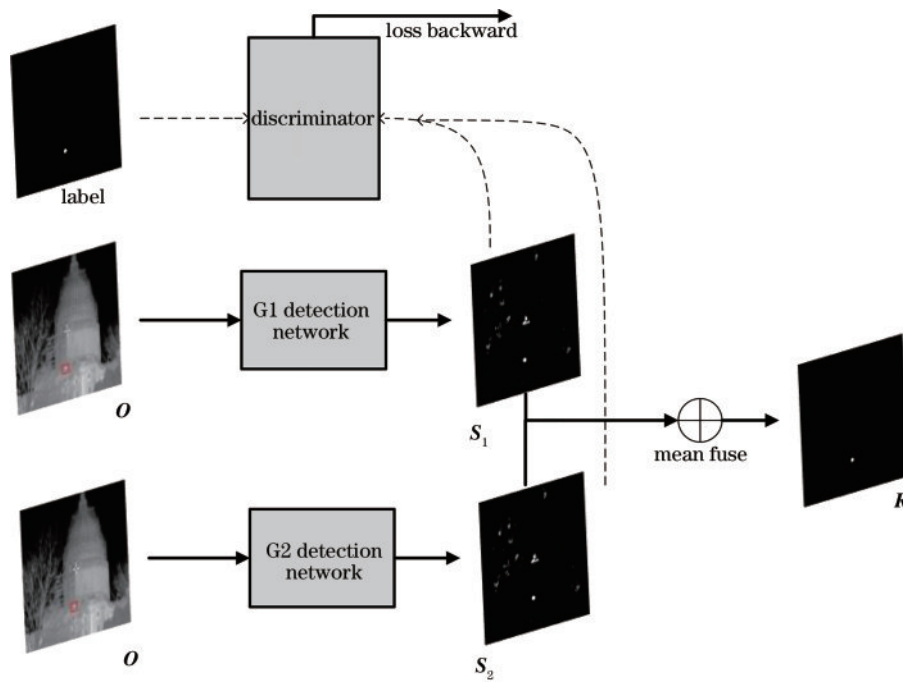


图 3 MDvsFA_cGAN 检测方案
Fig. 3 Detection scheme of MDvsFA_cGAN

使用条件生成对抗网络思想设计的红外小目标检测算法, G1 模块训练关注点在于提高对目标的检测度, G2 模块将背景抑制作为关注点, 融合输出结果 R 。

红外图像存在噪声和边缘干扰, 并且弱小目标在图中占比极小。图 2 算法缺乏强特征提取能力, 往往需要依靠后续处理对噪声和边缘进行过滤, 同时由于感受野较小, 缺乏全局特征, 因此小目标检测也存在难点。图 3 算法设计过于复杂, 网络整体耗时较长, 同时生成对抗网络的训练不稳定, 较难收敛。所提算法通过空洞卷积及跨层特征融合的方式, 融合高层语义信息和浅层空间信息, 结合较大感受野, 使得网络具有强特征提取能力, 并且引入注意力机制, 提高网络对噪声和边缘的过滤。

2.2 网络结构

作为一种端到端的图像分割算法, FCN 取消了传

统目标检测中的滑动窗口选取候选框, 舍弃了全连接层。因此, FCN 可以接收任意尺寸的输入图片, 编码过程中池化使得特征图缩小, 解码过程通过反卷积将特征图还原为输入尺寸, 从而实现逐像素点分类。

FCN 频繁的下采样操作会降低图片分辨率、丢失信息。因此在网络设计中, 舍弃下采样操作、使用空洞卷积替代普通卷积。空洞卷积相对于普通卷积, 特征图的像素点对应原图的感受区域更大, 能更好地融合局部特征和全局特征信息。从实验中可发现, 目标轮廓检测也优于其他算法。同时为了降低信息损失, 在每个空洞卷积层上设置等值填充, 保证分辨率和信息损失降到最低, 网络结构设计如图 4 所示。

由图 4 可知, 所提算法分为 3 个部分: 编码层、解码层和混合域模块。图 4 中 CBL 表示卷积层、batch normalization (BN) 层和 Leaky ReLU 层的组合。整个

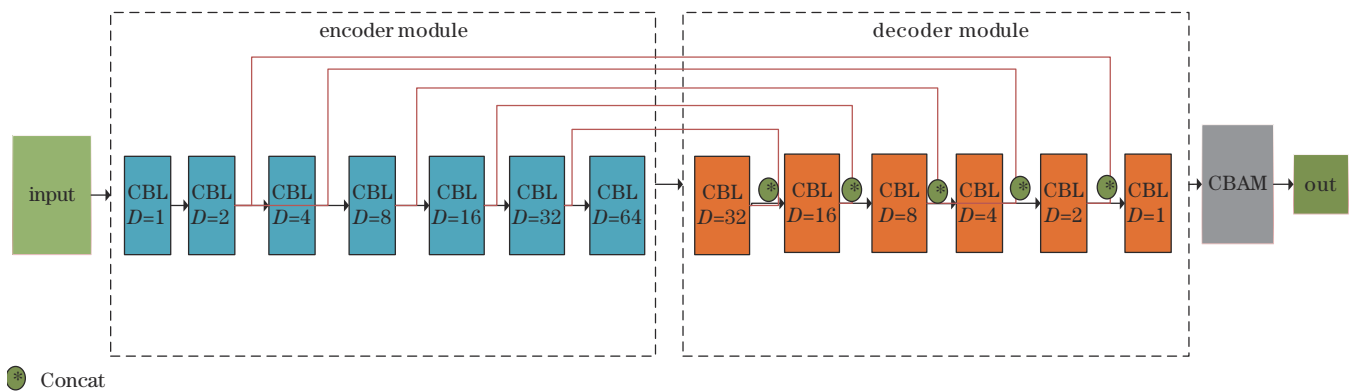


图 4 所提算法网络结构

Fig. 4 Network structure of proposed algorithm

网络的特征提取模块均为空洞卷积层,空洞扩张率 D 在编码层依次为 1、2、4、8、16、32、64,解码层空洞扩张率依次缩小,为 32、16、8、4、2、1。网络层的设计呈对称结构,逐层跨越拼接进行特征融合,充分融合网络深层的强语义信息和浅层的强空间位置信息,提高网络对目标的检测分割能力。在此基础上,利用混合域模

块的视觉显著性处理特征信息,进一步抑制强噪声点,提高目标识别度。

表 1 为编码层和解码层的设计细则,表中 CBL_X 为 CBL 的组合, X 表示网络层序号。Leaky ReLU 激活函数中的系数固定为 0.2,这种对称设计方式有益于增强网络鲁棒性。

表 1 编码-解码设计细则
Table 1 Design parameters of encoding-decoding

Combination	Layer	Kernel_size	Stride	Padding	Dilation	In_channel	Out_channel	ReLU_param
Encoder	CBL_1	3	1	1	1	3	64	0.2
	CBL_2	3	1	2	2	64	64	0.2
	CBL_3	3	1	4	4	64	64	0.2
	CBL_4	3	1	8	8	64	64	0.2
	CBL_5	3	1	16	16	64	64	0.2
	CBL_6	3	1	32	32	64	64	0.2
	CBL_7	3	1	64	64	64	64	0.2
Decoder	CBL_8	3	1	32	32	64	64	0.2
	CBL_9	3	1	16	16	128	64	0.2
	CBL_10	3	1	8	8	128	64	0.2
	CBL_11	3	1	4	4	128	64	0.2
	CBL_12	3	1	2	2	128	64	0.2
	CBL_13	3	1	1	1	128	64	0.2

2.3 混合域模块

混合域模块 CBAM 由通道注意力模块 (CAM) 和

空间注意力模块 (SAM) 组成,其整体流程如图 5 所示。

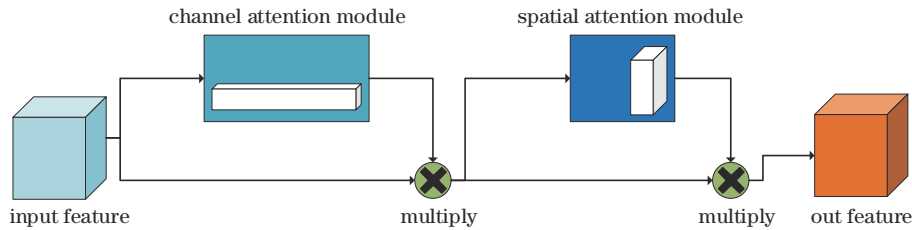


图 5 CBAM 整体流程
Fig. 5 Overall flow of CBAM

从图 5 可知,特征信息 F 经过分支进入 CAM,在通道维度上,通过增强可用特征的权重占比,降低干扰特征的权重,输出各通道权重 c 。将各通道权重 c 与特征信息 F 进行点乘,将结果 C_f 输入 SAM,提高重要特征区域的权重占比。将输出 s 与特征信息 C_f 点乘,输出最终上下文信息 S_f 。即

$$\begin{cases} c = \text{CAM}(F) \\ C_f = c \cdot F \end{cases}, \quad (3)$$

$$\begin{cases} s = \text{SAM}(C_f) \\ S_f = s \cdot C_f \end{cases}. \quad (4)$$

CAM 整体流程如图 6 所示。输入特征经过两条并行分支,分支包含池化层和感知机,分别为最大池化层和平均池化层。将这两部分的结果相加融合,输入 Sigmoid 激活层 δ ,从而计算每个通道的权重值 c_i 。即

$$c = \delta \left\{ \text{FC_layer} \left[\max_pool(O) \right] + \text{FC_layer} \left[\text{avg_pool}(O) \right] \right\}. \quad (5)$$

SAM 整体流程如图 7 所示。首先,利用 CAM 输出 c 结合特征信息 O 得到 C_f 。其次,将其作为 SAM 的输入特征,分别经过 max value reshape 和 avg value reshape,保持维度不变求取最大值和均值。接着,经过感知机变换映射,并将各分支输出拼接,经过卷积层 Conv 融合。最后,输入 Sigmoid 激活层 δ ,计算各空间区域的权重值 s_i 。即

$$s = \delta \left\{ \text{FC_layer} \left[\max_value(C_f) \right] + \text{FC_layer} \left[\text{avg_value}(C_f) \right] \right\}. \quad (6)$$

CBAM 通过空间和通道维度的注意力机制,将视觉显著性引入网络,类似人的视觉系统,有助于模型得

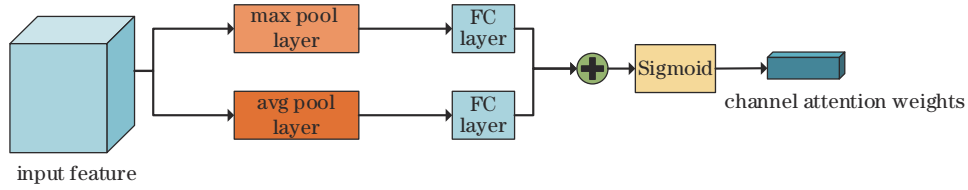


图 6 CAM 整体流程
Fig. 6 Overall flow of CAM

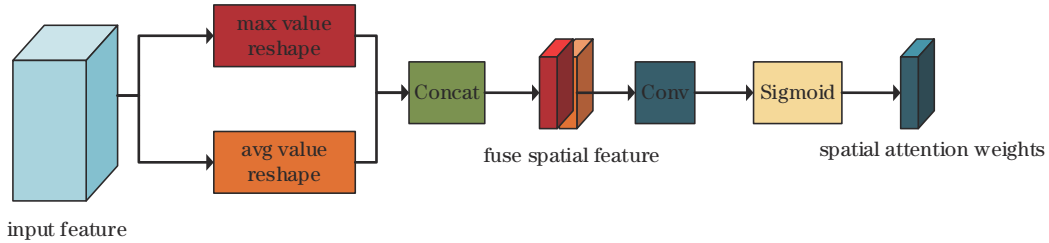


图 7 SAM 整体流程
Fig. 7 Overall flow of SAM

到目标增强且抑制背景的结果。

3 实验与分析

3.1 数据集

图 8 展示了真实红外图像及二值化标签。由于红外探测技术在天空背景下主要作为军事技术使用的客观原因限制,数据集存在一定敏感性,所以缺乏公开的红外小目标数据集。

本次实验所使用的数据集是在已搜寻得到的红外图像基础上,通过使用业界常用的扩充方式——光学系统的点扩散函数(PSF)^[25]获取的合成红外图像,从

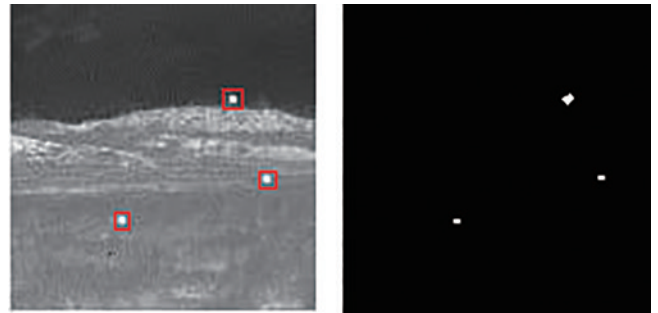


图 8 真实红外图像及二值化标签
Fig. 8 Real infrared images and binarised labels

而扩充数据集。图 9 显示了在掩膜上生成的点源。

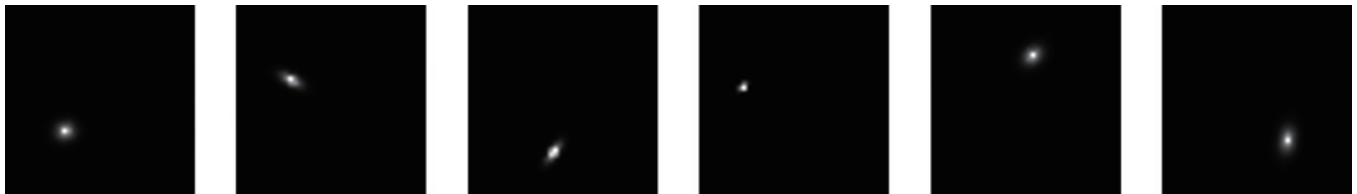


图 9 PSF 结合旋转变换生成点源
Fig. 9 Point source generated by PSF combined with rotation transformation

当距离所要观测的目标距离较大时,该目标释放的红外波段能量在红外探测器成像上接近一个点源。因此,可用 PSF 模拟生成这种点源,从而达到数据扩充的目的。图 10 对合成红外图像及标签进行了展示,整体趋近真实红外图像。合成的点源位置可描述为

$$I(x, y) = a \cdot \exp \left\{ -\frac{1}{2} \left[\frac{(x-x_0)^2}{\delta_x^2} + \frac{(y-y_0)^2}{\delta_y^2} \right] \right\}, \quad (7)$$

式中: δ_x, δ_y 分别表示水平、竖直方向发散程度,取值范围为 $[0.2, 1.25]$; (x, y) 为图像中任意坐标; (x_0, y_0) 是

点源中心,取值范围为 $(0 < x_0 + \delta_x < x, 0 < y_0 + \delta_y < y)$; a 是目标灰度的最大值,通常设置为 1。引入旋转角度变换,使合成点源具有多样性。

除此之外,针对合成点源方式背景单一、泛化性不强的问题,使用 Mosaic^[26] 数据增强技术进行数据集扩充。首先设置要裁剪的 C_x^{cut} 和 C_y^{cut} , 即

$$\begin{aligned} C_x^{\text{cut}} &= \text{random}(\partial \times w, \beta \times w); \\ C_y^{\text{cut}} &= \text{random}(\partial \times h, \beta \times h), \end{aligned} \quad (8)$$

式中: w 和 h 分别表示图像的宽高; ∂ 和 β 表示比例系数,设置 ∂ 为 0.2, β 为 0.8。图 11 展示了 4 张原图像经过 Mosaic 数据增强生成的合成图像。

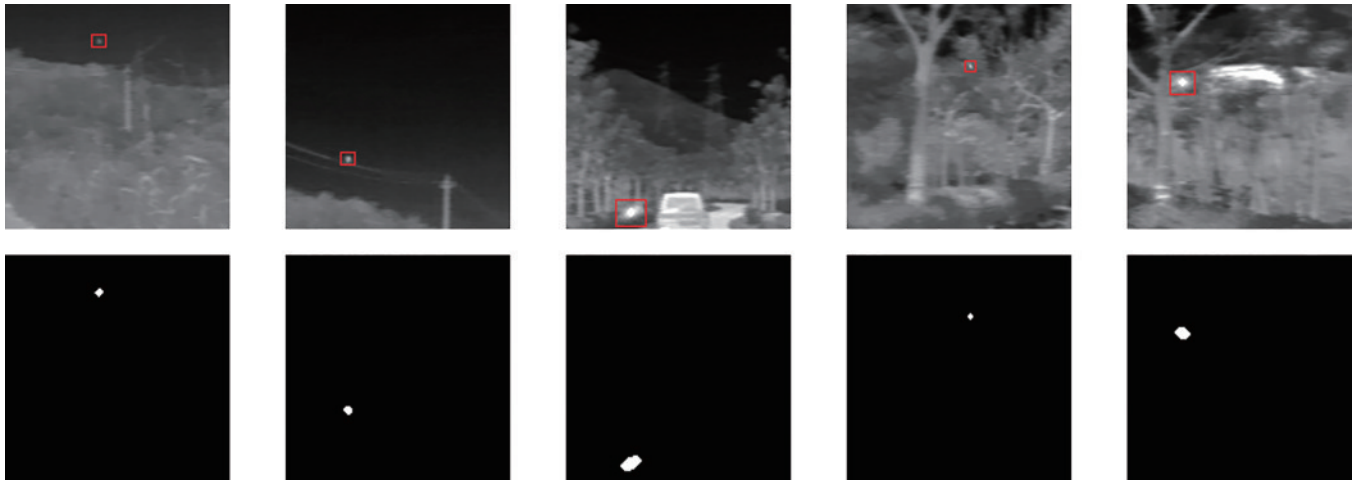


图 10 PSF 合成红外图像及标签

Fig. 10 Synthetic infrared images and labels based on PSF

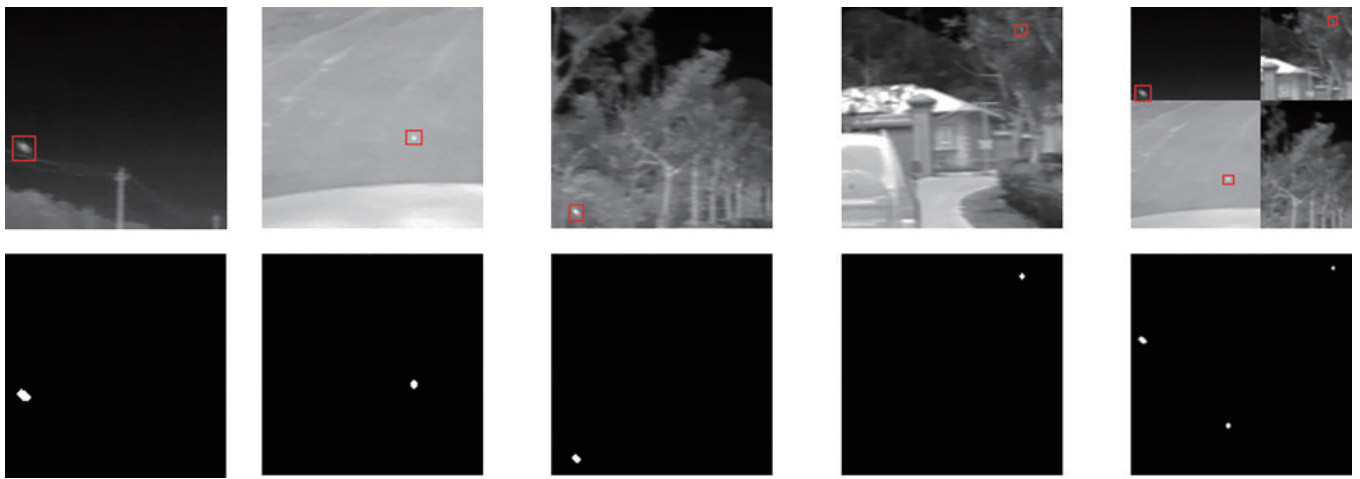


图 11 Mosaic 合成红外图像及标签

Fig. 11 Synthetic infrared images and labels based on Mosaic

为衡量数据集中目标与背景的灰度大小,计算图像中弱小目标的信噪比(SNR),即图像局部信噪比(LSNR):

$$R_{LSNR} = 10 \times \ln(E_T - E_B) / \delta_B, \quad (9)$$

式中: E_T 表示目标区域的均值; E_B 为背景区域均值; δ_B 表示背景区域的标准差。LSNR 用来描述目标信号的强弱,值越小表示信号越弱。国际光学工程学会将

LSNR 小于 5 dB 的目标定义为弱小目标。

通过以上扩充方法,得到 IR_GMS Dataset,其与公开数据集 IR_GS Dataset 的数量以及扩充方式如表 2 所示。IR_GS Dataset 和 IR_GMS Dataset 的验证集由真实红外图像构成,训练集由合成红外图像构成。使用 Mosaic 方法可以防止模型由于训练背景单一,而出现泛化性低的风险。

表 2 数据集扩充方式及细则

Table 2 Dataset expansion mode and details

Dataset	Synthesis	Mosaic augmentation	Total	Training set	Valid set	LSNR /dB
IR_GS Dataset	✓		10000	9900	100	2.83
IR_GMS Dataset	✓	✓	15000	14950	150	2.96

3.2 实验设置

利用 IR_GS Dataset 和 IR_GMS Dataset 进行模型的训练及验证,模型迭代轮数(epoch)设置为 30,优化器使用 Adam,学习率设置为 1×10^{-4} ,受限于 GPU 显存,batch_size 设置为 10。模型使用的硬件平台为 RTX 3070、英特尔 i5-10600kf,软件平台为

Windows 10。

构建模型损失函数,分别为 precision improved 损失函数(L_{PI})和 false alarm 损失函数(L_{FA})。计算公式如下:

$$L_{PI} = \frac{1}{n} (F_{pred} - F_{target})^2 \times F_{target}, \quad (10)$$

$$L_{FA} = \frac{1}{n} \sum (F_{pred} - F_{target})^2 \times (1 - F_{target}), \quad (11)$$

$$L = \alpha \times L_{PI} + \beta \times L_{FA}, \quad (12)$$

式中: F_{pred} 表示模型最终输出; F_{target} 为对应输入的真实标签; L_{PI} 负责计算 F_{pred} 与 F_{target} 对比所存在目标漏检的比值; L_{FA} 计算 F_{pred} 与 F_{target} 相比背景虚警的比值; α 和 β 是可学习调节参数, 初始值分别设置为 10 和 1。

为了与对比模型进行综合评估对比, 使用检测率 ($R_{precision}$)、召回率 (R_{recall})、虚警率 (R_{FPR}) 和 $F_{measure}$ ($R_{F_{measure}}$) 作为综合评估指标。图 12 中 TP 和 TN 代表预测正确的部分, 分别表示正确预测目标和正确预测背景; FP 和 FN 为模型预测错误的部分, 分别表示错误预测目标和错误预测背景。4 个评价指标的计算公式如下:

$$R_{precision} = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (13)$$

$$R_{recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (14)$$

$$R_{FPR} = \frac{N_{FP}}{N_{FP} + N_{TN}}, \quad (15)$$

$$R_{F_{measure}} = \frac{2R_{precision} \times R_{recall}}{R_{precision} + R_{recall}}, \quad (16)$$

式中: $R_{precision}$ 表示预测正确的目标占所有预测目标的比值; R_{recall} 表示预测正确的目标占实际目标的比值; R_{FPR} 表示预测错误的目标占所有预测背景的比值; $R_{F_{measure}}$ 作为 $R_{precision}$ 和 R_{recall} 的一种综合折中, 更能评估模型的性能; N_{TP} 表示正确预测目标的数量; N_{FP} 表示错误预测目标的数量; N_{FN} 表示错误预测背景的数量; N_{TN} 表示正确预测背景的数量。

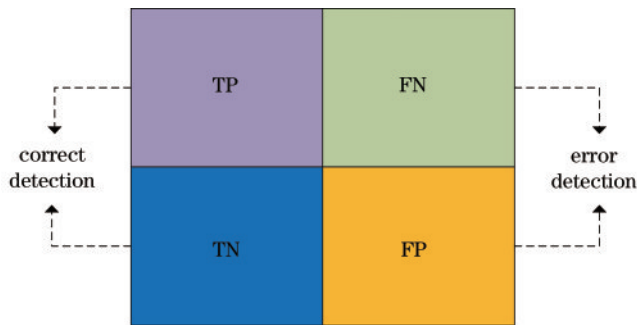


图 12 指标变量对应关系

Fig. 12 Correspondence of indicator variable

3.3 实验结果

所提算法在 30 个 epoch 的提升过程如图 13 所示。选取 8 张典型的红外图片, 查看网络在对应 epoch 的预测结果。

这 8 张红外图像包含了天空视角、遥感视角、天地视角以及混合背景视角, 背景较为复杂。图 13 选取模型在 5、10、20、30 个 epoch 的预测结果, 对比不同 epoch 下模型的拟合情况。模型在早期 epoch 上存在部分虚

警, 但是随着迭代的进行, R_{FPR} 逐渐下降, 并且 epoch 为 20 的预测图已基本等同于 epoch 为 30 的预测图, 可知模型已基本收敛。通过与 label 对比可知, 模型检测和虚警控制得较好。

在设置同等超参数及相同实验条件下, 各模型在 IR_GS Dataset 和 IR_GMS Dataset 的训练指标如表 3 所示。对比模型包括 MDvsFA_cGAN、DeepLab_ResNet、Fcn8x_ResNet、SegNet、UNet 和 UNet++。其中, Fcn8x_resNet 是文献[27]提出的红外小目标检测算法, DeepLab_ResNet、SegNet、UNet 和 UNet++ 是替换 Fcn8x_ResNet 算法的主干网络而得到的。为了验证 CBAM 对网络性能的提升, 将不含 CBAM 的算法命名为 dilation, 与所提算法 dilation_CBAM 做一个自身对比。

实验评估指标分别为 R_{FPR} 、 $R_{precision}$ 、 R_{recall} 、 $R_{F_{measure}}$ 以及实时性 (Speedup)。其中, Speedup 指标选取运行耗时最小的网络作为基准, 表示为 1X, 其他网络的实时性为该网络耗时与基准的比值。在两个数据集上: UNet++ 的 R_{FPR} 指标都取得了最好的效果, 其背景抑制最好; 而 dilation_CBAM 与 dilation 的其他指标都取得了最好的效果。在 IR_GS Dataset 上, dilation_CBAM 的 $R_{F_{measure}}$ 指标相比次优算法领先 4 个百分点, 在 IR_GMS Dataset 上也领先了 1.8 个百分点。得益于浅层网络设计, dilation 相较其他算法 Speedup 指标最好。将 dilation_CBAM 与 dilation 相比可发现, CBAM 的引入不仅提升了 $R_{precision}$, 而且在 R_{FPR} 指标上也有所降低, 验证了 CBAM 对背景抑制的增强作用。对于 R_{FPR} 指标, dilation_CBAM 与 UNet++ 差距较小, 仍具有实用性。对于其他指标, dilation_CBAM 相比 Unet++ 均有较大提升。

通过读取 $R_{F_{measure}}$ 值, 可以查看随着不同 epoch 各个模型的性能变化, 具体如图 14 所示。图 14(a)、(b) 分别表示 IR_GS Dataset、IR_GMS Dataset 上各模型的 $R_{F_{measure}}$ 指标。

由图 14 可知: 使用 Mosaic 数据增强技术的 IR_GMS Dataset, 各模型的 $R_{F_{measure}}$ 指标优于仅使用 PSF 的 IR_GS Dataset 上自身模型的表现, 证明 Mosaic 数据增强可以通过丰富图像背景和单张红外图像的小目标的方式, 提高数据集中红外图像的复杂度, 使模型在真实红外图像上更具有鲁棒性。通过图 14 还发现: 未加入混合域模块的 dilation 模型, 其网络波动性过大, 说明混合域模块通过视觉显著性方式过滤输出的特征信息, 在进行梯度反向更新时, 损失波动较小, 使算法具有稳定性。在 IR_GS Dataset 和 IR_GMS Dataset 上, 所提模型 dilation_CBAM 的表现均更好, 并且相较其他模型能快速收敛。

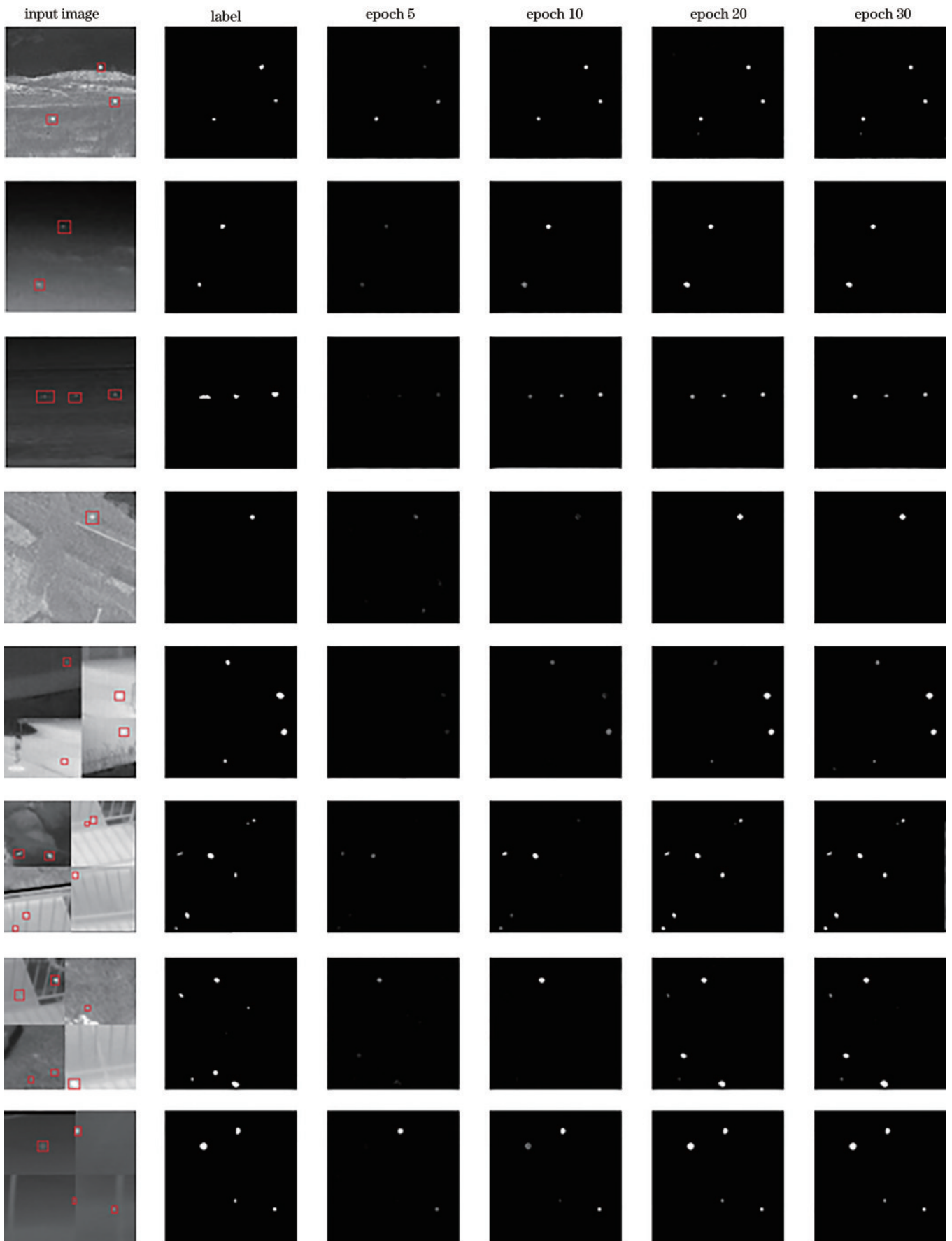


图 13 不同 epoch 效果比对

Fig. 13 Effects comparison of different epoch

表 3 典型深度学习模型实验结果对比

Table 3 Comparison of experimental results of typical deep learning model

unit: %

Dataset	Network	R_{FPR}	$R_{precision}$	R_{recall}	$R_{F_measure}$	Speedup
IR_GS Dataset	MDvsFA_cGAN	1.48×10^{-3}	0.740	0.380	0.502	1.6X
	DeepLab_ResNet	7.32×10^{-4}	0.240	0.189	0.212	5.1X
	Fcn8x_ResNet	4.23×10^{-4}	0.380	0.282	0.324	2.8X
	SegNet	4.36×10^{-4}	0.581	0.350	0.436	1.6X
	UNet	3.21×10^{-4}	0.526	0.390	0.448	1.2X
	UNet++	3.16×10^{-4}	0.624	0.414	0.498	3.5X
	dilation	5.64×10^{-4}	0.723	0.403	0.518	1X
dilation_CBAM	4.53×10^{-4}	0.747	0.445	0.558	1.2X	
IR_GMS Dataset	MDvsFA_cGAN	1.04×10^{-3}	0.683	0.500	0.578	1.6X
	DeepLab_ResNet	4.44×10^{-4}	0.330	0.342	0.336	5.1X
	Fcn8x_ResNet	5.01×10^{-4}	0.452	0.286	0.350	2.8X
	SegNet	3.49×10^{-4}	0.607	0.489	0.542	1.6X
	UNet	3.86×10^{-4}	0.597	0.465	0.522	1.2X
	UNet++	3.17×10^{-4}	0.584	0.483	0.529	3.5X
	dilation	5.46×10^{-4}	0.693	0.542	0.608	1X
dilation_CBAM	4.71×10^{-4}	0.670	0.588	0.626	1.2X	

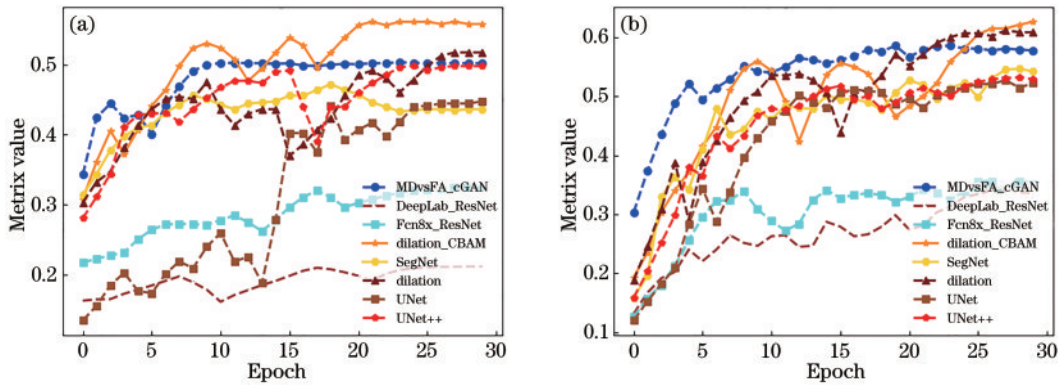


图 14 典型深度学习模型 $R_{F_measure}$ 指标对比。(a) IR_GS Dataset 数据集上各深度学习模型的 $R_{F_measure}$ 指标迭代情况; (b) IR_GMS Dataset 数据集上各深度学习模型的 $R_{F_measure}$ 指标迭代情况

Fig. 14 $R_{F_measure}$ index comparison of typical deep learning model. (a) $R_{F_measure}$ metric iterations for each deep learning model on the IR_GS Dataset; (b) $R_{F_measure}$ metric iterations for each deep learning model on the IR_GMS Dataset

4 对比测试

4.1 典型图像算法检测对比

为了直观评判各模型的预测差异,选取 6 张典型真实红外图像,包含山地、天空和建筑物等复杂且差异较大的背景,如图 15 所示。这 6 张图像从左至右依次编号为 1~6,可发现编号 2、3、6 图像中存在多个红外弱小目标,1、2、5 是单个红外弱小目标,5 中存在大量尖锐噪声,其余图像均包含复杂背景信息。通过图 15 可发现,综合表现最优为所提 dilation_CBAM 模型。

在处理背景较为单一和尖锐噪声较少的红外图像时,传统红外小目标检测算法基本能检测出目标;但在伴随较多尖锐噪声时,传统红外小目标检测算法的检测能力不太理想,检测结果存在大量误检和漏检现象。

深度学习模型能很好过滤尖锐噪声和强边缘信息,传统红外小目标检测算法易受噪声影响,导致检测结果与真实标签差异较大。

在不同背景下,dilation_CBAM 基本都检测出了弱小目标,并且背景抑制较好,预测结果虚警较低,而 dilation 相较 dilation_CBAM 存在漏检及虚警较高的问题。综合对比可知:这两个模型均优于此次实验的其他模型;其他模型部分虚警控制较好,但是存在漏检现象;部分能基本检测出弱小目标,但是虚警控制不太理想。综合考虑各个指标,所提 dilation_CBAM 模型在此次实验对比中表现最为理想。

4.2 公开红外图像测试

为进一步验证算法的泛化性,使用空天背景下的多序列红外小目标数据集^[28]进行实验。该数据集使用

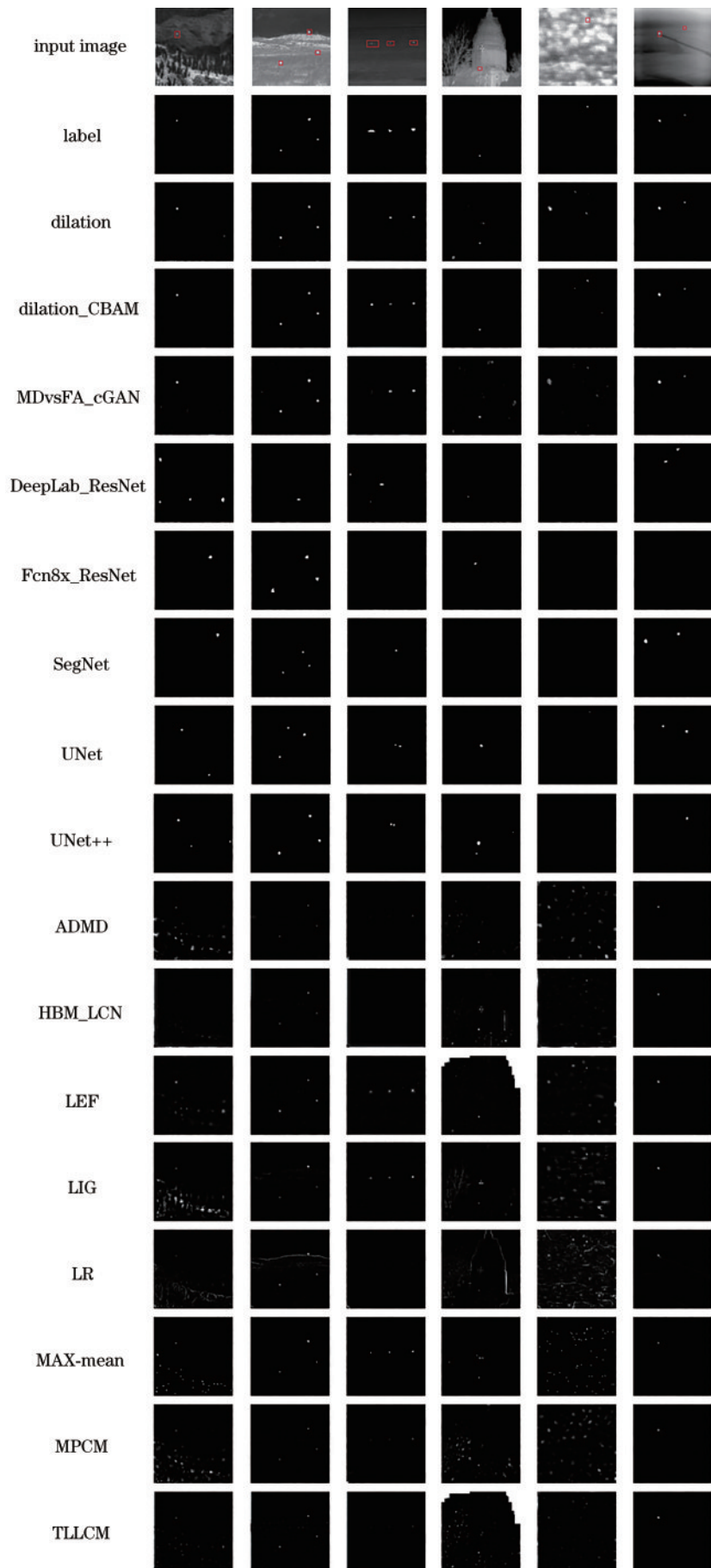


图 15 不同算法检测结果对比

Fig. 15 Comparison of detection results for different algorithms

一个或多个固定翼无人机目标提供数据,采集场景包括天空背景和复杂的野外背景(该数据集仅提供目标中心位置,未提供目标标注)。图 16 给出了测试图像

中典型的 12 张红外图像及算法预测结果。为了便于查看,图像中目标用蓝框圈出,预测得到的真实目标用红框区分,红外图像与其预测结果上下对应。

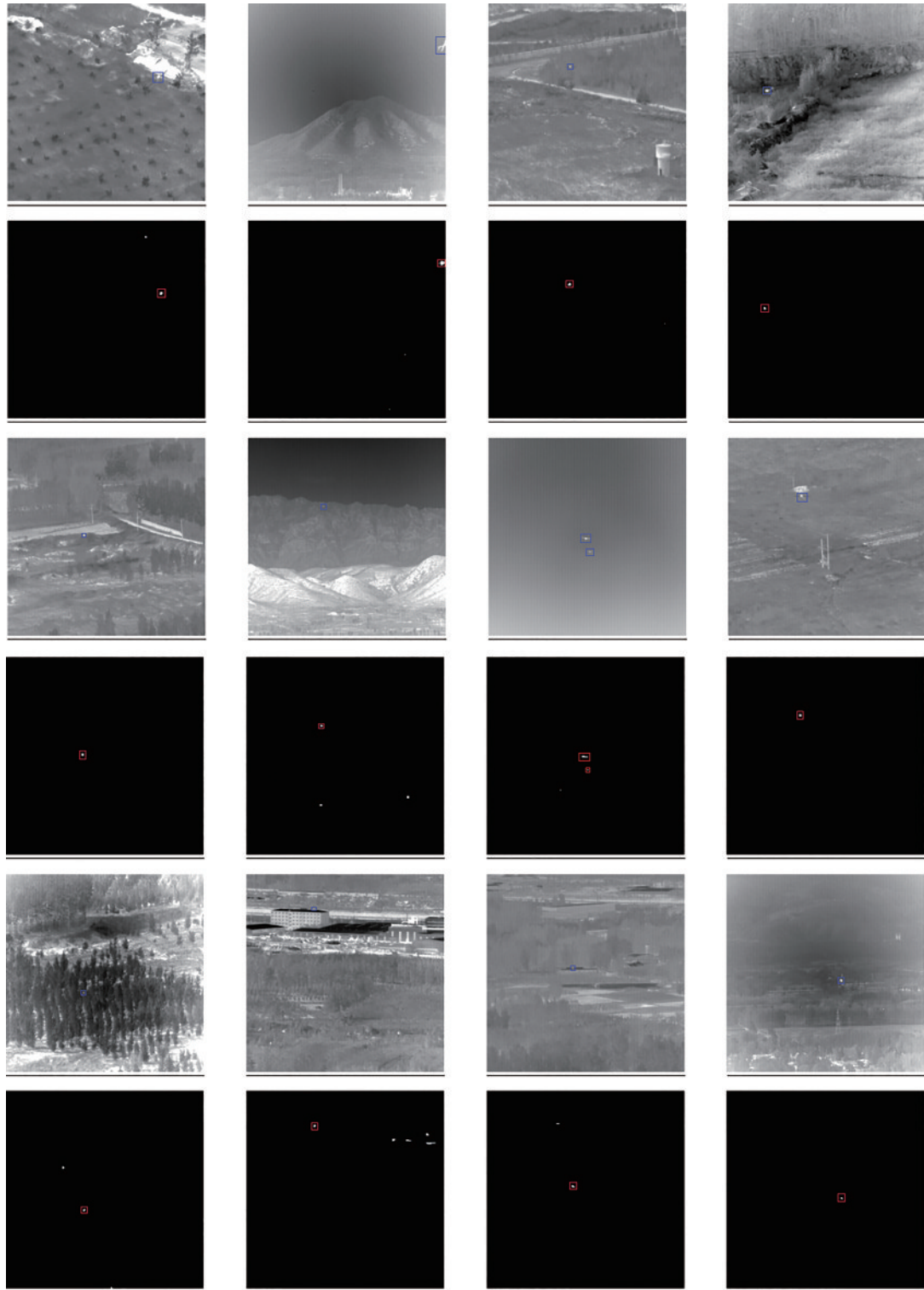


图 16 给定红外图像及算法预测结果

Fig.16 Given infrared images and predicted results

由图 16 可发现:dilation_CBAM 作为单帧红外弱小目标检测算法,在不依赖多序列提供的特征且网络

输入尺寸为 128 的前提下,所提算法在测试数据中的各个代表性场景,均能检测到真实目标。

5 结 论

将主干网络设计为编码-解码结构,并使用空洞卷积提高网络层的感受野,引入包含通道注意力机制和空间注意力机制的混合域 CBAM。通过对称的网络设计结合注意力机制的视觉显著性特点,实现对红外图像的弱小目标检测并且抑制虚警。实验结果表明,所提算法在 $R_{\text{precision}}$ 、 R_{FPR} 以及 Speedup 等指标上,其综合性能优于现有的典型红外小目标检测算法,且在不同背景下的红外图像上均能取得较好效果,鲁棒性较好。结合现有光学系统的 PSF,使用 Mosaic 数据增强技术对红外图像的扩充方式,为红外图像仿真数据提供了新的思路。未来将继续对网络结构进行优化,并尝试将网络结合传统红外小目标检测算法的结构做进一步深入研究。

参 考 文 献

- [1] 韩金辉. 基于人类视觉特性的复杂背景红外小目标检测研究[D]. 武汉: 华中科技大学, 2016.
Han J H. Infrared small target detection under complex background based on human visual system[D]. Wuhan: Huazhong University of Science and Technology, 2016.
- [2] 解婷, 陈忠, 马荣毅. 一种基于 PGF、BEMD 和局部逆熵的新型红外小目标检测方法[J]. 红外与毫米波学报, 2017, 36(1): 92-101.
Xie T, Chen Z, Ma R Y. A novel method for infrared small target detection based on PGF, BEMD and LIE[J]. Journal of Infrared and Millimeter Waves, 2017, 36(1): 92-101.
- [3] Chen C L P, Li H, Wei Y T, et al. A local contrast method for small infrared target detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 574-581.
- [4] Moradi S, Moallem P, Sabahi M F. Fast and robust small infrared target detection using absolute directional mean difference algorithm[J]. Signal Processing, 2020, 177: 107727.
- [5] Nie J Y, Qu S C, Wei Y T, et al. An infrared small target detection method based on multiscale local homogeneity measure[J]. Infrared Physics & Technology, 2018, 90: 186-194.
- [6] Wei Y T, You X G, Li H. Multiscale patch-based contrast measure for small infrared target detection[J]. Pattern Recognition, 2016, 58: 216-226.
- [7] Zhang H, Zhang L, Yuan D, et al. Infrared small target detection based on local intensity and gradient properties[J]. Infrared Physics & Technology, 2018, 89: 88-96.
- [8] 何玉杰, 李敏, 张金利, 等. 基于相关滤波器的红外弱小目标检测算法[J]. 光学学报, 2016, 36(5): 0512001.
He Y J, Li M, Zhang J L, et al. Infrared small target detection method based on correlation filter[J]. Acta Optica Sinica, 2016, 36(5): 0512001.
- [9] Xu Y H. Real-time detection algorithm for small space targets based on max-median filter[J]. Journal of Information and Computational Science, 2014, 11(4): 1047-1055.
- [10] 吴巍, 彭嘉雄, 王海晖. 红外图像序列小目标的特性分析与检测[J]. 红外与激光工程, 2002, 31(2): 146-149.
Wu W, Peng J X, Wang H H. Characteristics analysis and detection of small target in infrared image sequences[J]. Infrared and Laser Engineering, 2002, 31(2): 146-149.
- [11] 楼晨风, 张湧, 刘亚. 基于互补梯度增强的红外阵列扫描图像小目标检测[J]. 光学学报, 2021, 41(21): 2104001.
Lou C F, Zhang Y, Liu Y. Small target detection of infrared linear array image based on complemented gradient enhancement[J]. Acta Optica Sinica, 2021, 41(21): 2104001.
- [12] Wan M J, Gu G H, Qian W X, et al. Total variation regularization term-based low-rank and sparse matrix representation model for infrared moving target tracking[J]. Remote Sensing, 2018, 10(4): 510.
- [13] Gao C Q, Meng D Y, Yang Y, et al. Infrared patch-image model for small target detection in a single image[J]. IEEE Transactions on Image Processing, 2013, 22(12): 4996-5009.
- [14] Zheng C Y, Li H. Small infrared target detection based on low-rank and sparse matrix decomposition[J]. Applied Mechanics and Materials, 2012, 239/240: 214-218.
- [15] 王勃凡, 赵海涛. 基于径向基激活函数的高光谱小目标检测[J]. 光学学报, 2021, 41(23): 2311001.
Wang B F, Zhao H T. Small object detection in hyperspectral images based on radial basis activation function[J]. Acta Optica Sinica, 2021, 41(23): 2311001.
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3431-3440.
- [17] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [18] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [19] Zhou Z W, Siddiquee M M R, Tajbakhsh N, et al. UNet++: a nested U-net architecture for medical image segmentation[EB/OL]. (2018-07-18)[2021-05-04]. <https://arxiv.org/abs/1807.10165>.
- [20] Wang H, Zhou L P, Wang L. Miss detection vs. false alarm: adversarial learning for small object segmentation in infrared images[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 8508-8517.
- [21] 荣生辉, 刘刚, 周慧鑫, 等. 基于改进的剪切波变换和引导滤波的红外弱小目标背景抑制方法[J]. 光子学报,

- 2015, 44(2): 0210002.
- Rong S H, Liu G, Zhou H X, et al. Infrared dim and small target background suppression based on the improved shearlet transform and the guide filter[J]. Acta Photonica Sinica, 2015, 44(2): 0210002.
- [22] 秦翰林, 黄洋, 姚柯柯, 等. 多尺度核局部归一化的红外图像背景抑制[J]. 强激光与粒子束, 2012, 24(5): 1063-1066.
- Qin H L, Huang Y, Yao K K, et al. Multi-scale kernel local normalization for infrared image background suppression[J]. High Power Laser and Particle Beams, 2012, 24(5): 1063-1066.
- [23] 侯春萍, 王霄聪, 夏晗, 等. 基于双通路生成对抗网络的红外与可见光图像融合方法[J]. 激光与光电子学进展, 2021, 58(14): 1410024.
- Hou C P, Wang X C, Xia H, et al. Infrared and visible image fusion method based on dual-channel generative adversarial network[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1410024.
- [24] 刘峰, 郭猛, 王向军. 基于跨尺度融合的卷积神经网络小目标检测[J]. 激光与光电子学进展, 2021, 58(6): 0610012.
- Liu F, Guo M, Wang X J. Small target detection based on cross-scale fusion convolution neural network[J]. Laser & Optoelectronics Progress, 2021, 58(6): 0610012.
- [25] 林两魁, 王少游, 唐忠兴. 基于深度卷积神经网络的红外过采样扫描图像点目标检测方法[J]. 红外与毫米波学报, 2018, 37(2): 219-226.
- Lin L K, Wang S Y, Tang Z X. Point target detection in infrared over-sampling scanning images using deep convolutional neural networks[J]. Journal of Infrared and Millimeter Waves, 2018, 37(2): 219-226.
- [26] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-05-04]. <https://arxiv.org/abs/2004.10934>.
- [27] 刘俊明, 孟卫华. 融合全卷积神经网络和视觉显著性的红外小目标检测[J]. 光子学报, 2020, 49(7): 0710003.
- Liu J M, Meng W H. Infrared small target detection based on fully convolutional neural network and visual saliency[J]. Acta Photonica Sinica, 2020, 49(7): 0710003.
- [28] 回丙伟, 宋志勇, 范红旗, 等. 地面/空中背景下红外图像微弱小飞机目标检测与跟踪数据集(V1)[EB/OL]. (2019-10-28)[2021-05-08]. <https://datapid.cn/31253.11.sciencedb.902>.
- Hui B W, Song Z Y, Fan H Q, et al. A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background(V1) [EB/OL]. (2019-10-28)[2021-05-08]. <https://datapid.cn/31253.11.sciencedb.902>.