

文章编号 1004-924X(2023)20-3034-16

空间信息自适应调控和特征对齐的 红外甲烷实例分割

何自芬, 曹辉柱, 张印辉*, 庄 宏
(昆明理工大学 机电工程学院, 云南 昆明 650000)

摘要:传统接触式甲烷泄漏传感器检测范围小且效率低,而结合非接触式红外热成像的机器视觉算法可实现远距离、大范围红外甲烷实例分割,对于提高甲烷检测效率及保障人员安全具有显著优势。然而远距离甲烷气体图像轮廓模糊、泄漏的甲烷气体与背景对比度较低且形状易受大气流动因素影响等问题限制了红外甲烷实例分割性能。针对上述问题,本文提出一种空间信息自适应调控和特征对齐的网络模型(Adaptive spatial information regulation and Feature alignment Network, AFNet)实现甲烷泄漏红外实例分割。首先,为增强模型的特征提取能力,提出自适应空间信息调控模块赋予主干网络不同尺度残差块自适应权重丰富模型提取的特征空间;其次,构建加权双向金字塔弥补特征金字塔自顶而下的特征传播方式导致的低层特征空间位置和实例边缘信息弥散丢失问题,以适应甲烷气体复杂轮廓变化下前景目标定位检测和轮廓分割需求。最后,设计原型特征对齐模块捕获长距离气体特征之间的语义关系丰富原型语义信息量以改善生成目标掩码质量提高甲烷气体分割精度。实验结果表明,本文提出的 AFNet 模型 AP50@95, AP50 定量分割精度分别达到 42.42%, 92.18%, 相比于原始 Yolact 模型分割精度,分别提高 9.79%, 6.18%, 推理速度达到 36.80 frame/s, 满足甲烷泄漏分割需求。实验结果验证了本文算法对红外甲烷泄漏分割的有效性和工程实用性。

关键词:红外甲烷;自适应调控;特征对齐;特征金字塔;实例分割

中图分类号:TP394.1;TH691.9 文献标识码:A doi:10.37188/OPE.20233120.3034

Spatial information adaptive regulation and feature alignment for infrared methane instance segmentation

HE Zifen, CAO Huizhu, ZHANG Yinhu*, ZHUANG Hong

(Faculty of Mechatronics and Electrical Engineering, Kunming University of Science and Technology,
Kunming 650000, China)

* Corresponding author, E-mail: yinhui_z@163.com

Abstract: Conventional contact methane leak sensors suffer from a small detection range and low efficiency, but machine vision algorithms combined with non-contact infrared thermal imaging can make infrared methane instance segmentation possible at long distances and large ranges. This is a significant advantage for improving methane detection efficiency and ensuring personnel safety. However, the segmentation performance of infrared methane instances is limited by such problems as blurred contour and low contrast between the leaking methane gas and the background, and it can be affected by atmospheric flow factors. In response to these problems, an adaptive spatial information regulation and feature alignment network (AF-

收稿日期:2023-04-03;修订日期:2023-05-08.

基金项目:国家自然科学基金资助项目(No. 62171206, No. 62061022)

Net) is proposed to segment infrared instances of methane leakage. First, to enhance the model's feature extraction, an adaptive spatial information regulation module is proposed to endow the backbone network with adaptive weights for different scale residual blocks, which enrich the feature space extracted by the model. Second, to meet the requirements of foreground target positioning detection and contour segmentation under complex methane gas contours, a weighted bidirectional pyramid is designed to reduce the diffusion, loss of spatial location, and instance edge information in low-level features, which are caused by the top-down propagation of the feature pyramid. Finally, a prototype feature alignment module is designed to capture the semantic relationships between long-distance gas features, enriching the semantic information of the prototype and improving the quality of generated target masks to improve the methane instance segmentation accuracy. Experimental results show that the proposed AFNet model achieves AP50@95 and AP50 quantitative segmentation accuracies of 42.42% and 92.18%, which are improved by 9.79% and 6.18% compared with the original Yolact, respectively. In addition, the inference speed achieves 36.80 frames/s and meets the requirements of methane leakage segmentation. The experimental results validate the effectiveness and engineering practicality of the algorithm proposed for infrared methane leakage segmentation.

Key words: infrared methane; adaptive regulation; feature alignment; feature pyramids; instance segmentation

1 引言

甲烷是天然气与煤气的主要成分,广泛存在于煤炭化工、石油开采、石油化工等化工行业,同时也是日常生活中主要清洁能源^[1]。另外,甲烷作为温室气体之一,其对全球变暖的影响远远超过二氧化碳^[2]。尽管甲烷在大气中寿命只有12年,但具有相较于二氧化碳28~34倍的蓄热能力,使得它成为全球变暖的主要来源^[3]。甲烷具有无色无味且易燃易爆的化学性质,若无检测设备其泄漏就很难被察觉。一旦接触到热源或者明火,就可能导致严重的安全事故,如火灾、爆炸和环境污染,对社会的生命和财产构成巨大威胁。因此,减少甲烷气体泄漏事故已成为化工生产中日益关注的问题,发展针对甲烷气体泄漏的高效安全检测和分割技术用于实时监测甲烷泄漏、快速定位甲烷泄漏源,以减少人民生命财产损失、保护环境,做到降本增效经济意义重大。

传统气体泄漏检测主要依靠接触式气体传感器对管道中泄漏气体进行检测,包括人工检测、气体采样、土壤检测等^[4-6]。此类方法不仅耗费大量的人力财力和时间,而且对管道、储气库及周边环境造成一定的破坏^[7],同时检测结果易受安装位置和外界环境风向等因素影响,无法达

到快速定位泄漏源的目的,且检测范围小、检测效率低。近年来,非接触式红外热成像探测技术,以其探测距离远、降低了检测难度和风险因素,同时还有灵敏度高、动态直观等优点,在石化装置及工艺管线的气体泄漏检测领域得到广泛应用^[8]。与可见光气体成像不同,红外成像利用有害气体与背景之间的温度差,将无法直接观察到的红外辐射转换为灰度图像,有效将有害气体与周围环境分开,从而实现对有害气体进行泄漏监测。红外热成像技术可以更加直观地检测出某种气体物质的存在,并且能够准确地确定其位置,可以帮助我们更好理解和识别特定气体。

由于红外热成像技术具有快速响应、非接触测量和快速定位气敏传感器难以感知到的微小泄漏源等优势,因此在气体检漏领域得到了广泛应用^[9]。尽管红外热像仪可以手持或安装在固定位置,但目前仍然依赖于人工进行视检使得人力成本增加^[10]。近年来,由于人工智能技术飞速发展,以卷积神经网络为基础的深度学习算法已成为计算机视觉领域的重要技术手段,并且正被广泛应用于视觉检测中以替代传统的人工巡检^[11]。

研究人员近年在红外气体泄漏检测方面投入了大量的工作,如基于目标气体红外图像进行

分类^[12-14]和目标检测^[15]取得了很好的效果,但此类检测算法无法将泄漏气体从背景中分割出来也无法给出泄漏气体的轮廓边界,不利于后续对各个点源的泄漏等级进行量化以便制定维修方案。

实例分割是目标检测与语义分割在模型层面的统一,要求正确地检测并分割图像中每个对象实例,实现了像素维度的图像目标轮廓识别。目前广泛使用的基于深度学习的实例分割方法可分为两类。一类是双阶段实例分割算法,如 Mask R-CNN^[16]算法,采用深度卷积神经网络提取特征,在图像实例分割领域取得了较好的结果,在此基础上 Mask Scoring R-CNN^[17], HTC^[18]等方法相继被提出。另一类是单阶段实例分割方法,如 Yolact 在 RetinaNet^[19]基础上,对预测头模块放弃引入特征定位的步骤且增加一个掩码分支达到实例分割的效果,是第一个实现实时实例分割的方法。与之类似的还有 SOLOv1-v2^[20-21]和 CondInst^[22]等方法。此类单阶段实例分割算法不需要生成候选目标,直接通过网络预测目标的位置和类别,在保证精度的同时具有更快的检测速度。

综上所述,为保证模型对红外甲烷气体准确分割,利用红外甲烷气体图像完成特征学习,并克服由于红外甲烷气体目标实例尺度变化及拓扑形变导致分割边缘模糊和特征弥散丢失等问题,实现红外甲烷泄漏有效分割。本文以 Yolact^[23]为基础模型,提出一种空间信息自适应调控和特征对齐的实例分割网络模型(Adaptive spatial information regulation and Feature alignment Network, AFNet),在速度和准确性之间实现了良好的平衡,改进方法与目的如下:

(1) 针对甲烷气体泄漏扩散区域形变复杂,模型采用 ResNet^[24]残差网络作为主干网络提取红外甲烷气体特征,通过自适应空间信息调控模块丰富主干网络中空间位置信息以及调控目标特征区域语义特征。

(2) 针对特征金字塔自顶而下单一特征传播路径和模型过深导致的目标气体特征空间细节信息弥散丢失问题构建加权双向金字塔^[25]模块丰富深层特征图下目标实例的边缘纹理信息,完善细粒度像素级回归任务以增强模型的目标区

域聚焦能力。

(3) 针对原型掩码分支利用全卷积网络^[26]网络进行特征提取时仅简单利用卷积层增大其感受野,忽略不同维度特征图像之间的特征差异与不同特征对于目标分割精度的重要性,提出特征对齐模块捕获长距离范围内像素点之间的语义关系,对齐不同特征空间特征信息以生成高质量掩码提高分割精度。

2 网络模型

2.1 网络整体结构

由于环境温度变化和气流扰动等外界因素导致甲烷气体分子微观上不停做无规则热运动,宏观上甲烷气体在扩散过程中扩散区域呈现出不规则形状变化,气体边缘轮廓始终呈现为模糊不清的混沌状态,而且甲烷气体在红外成像下存在着缺少色彩信息,空间和纹理细节信息匮乏等问题,本文以 Yolact 模型为基础,构建出一种精度高、掩码质量好的空间信息自适应调控和原型特征对齐的红外甲烷泄漏实例分割模型 AFNet。AFNet 模型主要由两个阶段组成,分别为特征提取阶段和特征后处理阶段。特征提取阶段由 ResNet 主干网络、加权双向金字塔网络和原型特征对齐模块组成;特征后处理阶段由预测头分支组成。在特征提取阶段,首先将红外甲烷气体图像输入 ResNet 主干网络中,通过自适应空间信息调控模块增强骨干网络对目标特征的自适应聚合能力,丰富模型提取特征的信息量和增强上下文相关性,并对分割结果贡献大的特征权重进行自适应校准,调整模型关注区域,减小背景干扰。其次,引入加权双向金字塔弥补特征金字塔自顶而下的特征传播方式导致的低层特征空间位置和实例边缘信息弥散丢失问题,以适应甲烷气体复杂轮廓变化下前景气体目标定位检测和轮廓分割需求。最后设计原型特征对齐模块捕获长距离范围内甲烷气体像素点之间的语义关系,对齐不同特征空间特征信息从而丰富原型语义信息量以生成高质量掩码提高分割精度。在特征后处理阶段,将加权双向金字塔的多尺度特征图作为预测头网络输入,经过分类分支头、边界框分支头和掩码分支头三个并行分支头得到输出结果。AFNet 模型整体结构如图 1 所示。

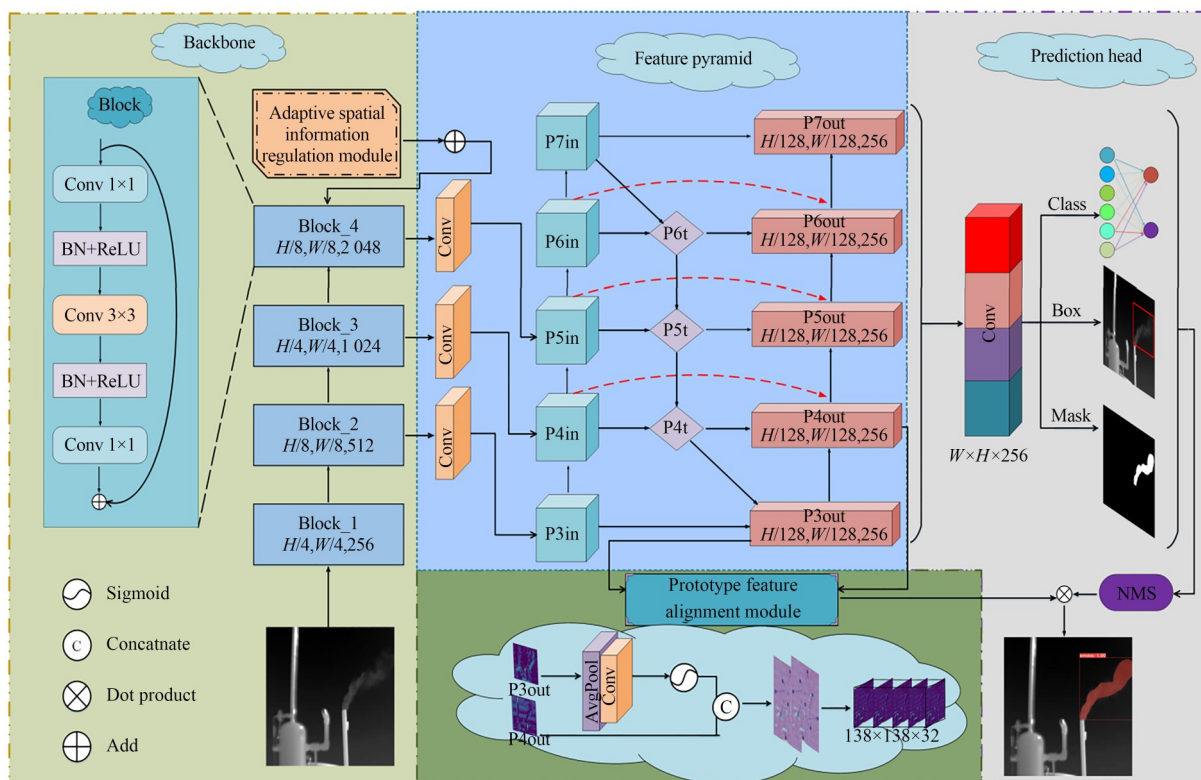


图 1 AFNet网络架构

Fig. 1 AFNet network architecture

2.2 自适应空间信息调控

由于红外甲烷气体图像存在缺少色彩、纹理信息,轮廓复杂多变且边缘模糊等问题,仅由主干网络浅层或深层学习到的特征往往难以表征气体全局特征。在特征提取网络阶段,ResNet主干网络分别由Block_1,Block_2,Block_3和Block_4个特征提取模块组成,4个模块特征图表征的语义信息依次由浅到深。浅层特征由于涉及的特征映射变换少,所以保留更多原始输入纹理特征。因此,

浅层特征更有助于定位目标实例的边缘纹理,从而更好地进行分割。而深层特征经过逐层抽象变换可以更好地捕捉目标实例的语义要素,以指导模型的最终分割。针对上述问题,本文提出自适应空间信息调控模块(Adaptive Spatial Information Regulation Module, ASIR),在ResNet主干网络中嵌入自适应空间信息调控模块丰富模型提取的特征空间有效表征甲烷气体特征,自适应空间信息调控模块结构如图2所示。

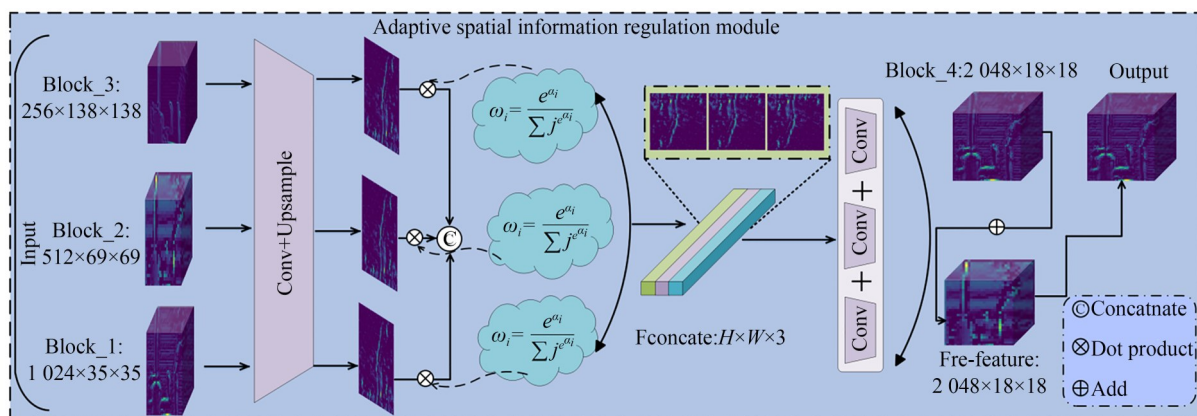


图 2 自适应空间信息调控模块

Fig. 2 Adaptive spatial information regulation module

自适应空间信息调控模块首先对主干网络 Block_1, Block_2 和 Block_3 分别采用卷积层进行通道信息压缩操作将三个特征图通道压缩至 1 维, 分别与自适应权重因子 ω_1, ω_2 和 ω_3 相乘。在模型训练过程中, 通过 PyTorch 深度学习运行库生成一组尺寸等于调控模块数量且初始值为一的全一张量, 然后通过公式(1)所示的归一化方式, 分别将每个张量值转化为所对应的指数形式的权重相对于所有权重的比例值, 从而得到初始权重因子 ω_i , 其中 $i = 1, 2, 3$ 。

根据 Block_1, Block_2 和 Block_3 特征层所包含空间信息量贡献程度大小差异, ω_1, ω_2 和 ω_3 在训练过程中依赖神经网络的反向传播进行梯度更新, 达到自适应调整目的。然后再将 Block_1, Block_2 和 Block_3 通道压缩后的特征图进行通道拼接得到具有自适应能力的特征 $F_{\text{concate}}(H \times W \times 3)$; 进一步将特征 $F_{\text{concate}}(H \times W \times 3)$ 经由卷积层使通道维度信息恢复到 2 048, 达到对特征空间信息压缩和激励目的。由于对特征空间信息进行了压缩与激励, 拼接后的特征张量每个像素点均包含了不同尺度下该位置所有通道信息, 将其送入堆叠的卷积层进行自适应特征学习, 有效实现不同尺度下特征张量之间空间信息的流动, 丰富模型提取特征细节信息量。最后将重编码后的特征 $F_{\text{re-feature}}(H \times W \times 2\,048)$ 与最具抽象语义信息的 Block_4 模块输出特征图进行像素点乘最终得到自适应多尺度空

间信息特征图 $F_{\text{adaptive}}(H \times W \times 2\,048)$, 实现低维空间信息在高级语义信息中自适应调控, 达到对空间信息权重的自适应校准。自适应空间信息调控模块如公式(1)所示:

$$\begin{aligned} F_{\text{concate}} &= \text{Conv}_{3 \times 3}(\text{Concate}(\omega_1 \times F_{\text{Block}_1} + \\ &\quad \omega_2 \times F_{\text{Block}_2} + \omega_3 \times F_{\text{Block}_3})) \\ F_{\text{adaptive}} &= F_{\text{concate}} \otimes F_{\text{Block}_4} \\ \omega_i &= \frac{e^{\alpha_i}}{\sum_j e^{\alpha_j}} (i = 1, 2, 3; j = 1, 2, 3), \end{aligned} \quad (1)$$

其中: ω_i 表示自适应权重, α_i 为初始化指数权重, α_j 为特征权重。 F_{concate} 表示具有自适应能力的融合特征, $\text{Conv}_{3 \times 3}$ 表示 3×3 卷积操作, F_{adaptive} 表示最终输出特征图, Concate 表示通道拼接操作, $\otimes$ 表示像素点乘。

训练时主干网络中四个残差块 Block_1, Block_2, Block_3 和 Block_4 特征图进行可视化如图 3 所示(彩图见期刊电子版), 第一行为原始 ResNet 主干网络特征图, 第二行为空间信息调控模块输出特征图。从图 3 中红色标注部分可以看出经过空间信息调控残差块特征图中气体目标轮廓特征更为明显, 边缘区分较好, 纹理细节信息丰富, 极大程度降低复杂背景对前景分割质量的影响, 对特征的拟合效果明显优于 Yolact 原始模型, 在 Block_2, Block_3, Block_4 深度递增特征图上特征学习更为突出且在 Block_4 特征图背景抑制效果明显。

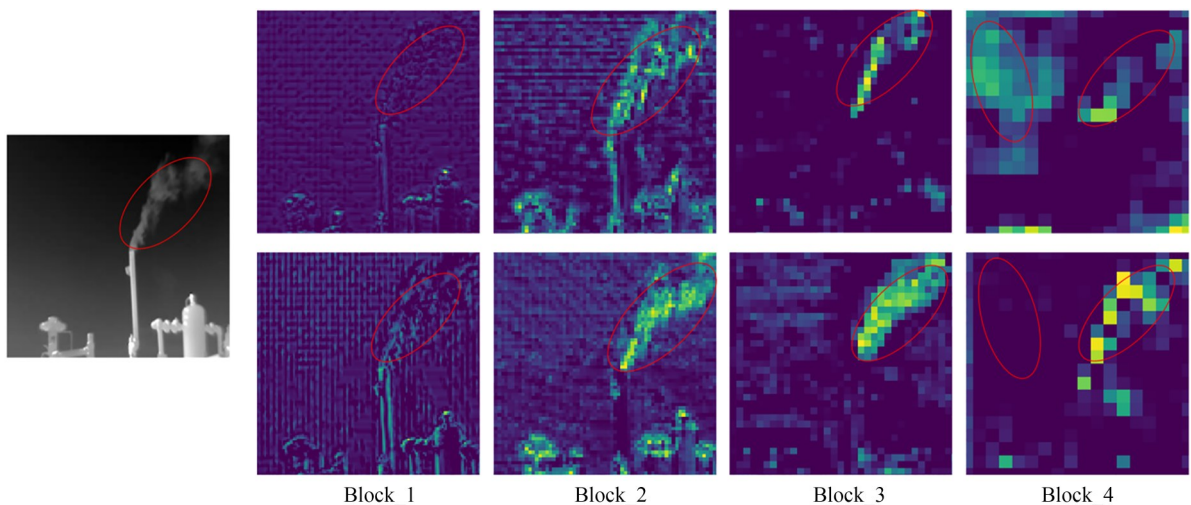


图 3 空间信息调控模块特征图对比

Fig. 3 Comparison of feature maps of spatial information regulation module

2.3 加权双向金字塔

Yolact 模型采用特征金字塔结构^[27]实现不同尺度特征图的信息融合对气体目标进行预测。特征金字塔通过自顶向下和特征横向映射融合多尺度特征,但该特征映射传播方式存在一定问题。由于红外甲烷气体图像具有尺度差异大、纹理细节信息稀缺,不同深度特征经过单一自顶向下的特征横向映射和大量卷积操作会使得气体目标特征细节信息丢失,不利于精确识别气体目标。针对上述问题 PANet^[28]在特征金字塔的基础上增加自底向上路径,将浅层信息直接与高层信息融合,减少各网络层聚合的特征信息经过骨干网络造成的信息丢失问题,从而达到丰富网络提取的特征空间和使定位信息更为准确的目的。PANet 增加的自底向上结构在一定程度上使模型提取的特征信息更为丰富,但其所加分支也带

来了附加参数量,使网络模型的时空复杂度增大,计算量有较大程度增加。BiFPN 即加权双向金字塔在保留 PANet 高精度的同时,对多尺度特征信息融合方式进行调整,针对 PANet 结构中部分节点只有一个输入,且没有进行特征融合操作的问题,在连接时将贡献度较低的输入节点删除以减少网络参数量使特征融合网络更加高效。同时为缓解网络链路较多造成特征损失问题,在输入节点和输出节点之间增加跳跃连接以实现特征的有效融合。因此本文引入加权双向金字塔减少不同深度特征横向映射过程中细节信息弥散丢失,丰富深层特征图中气体目标实例边缘纹理信息,完善细粒度像素级回归任务以适应甲烷气体复杂轮廓变化下轮廓分割需求,其结构如图 4 所示。

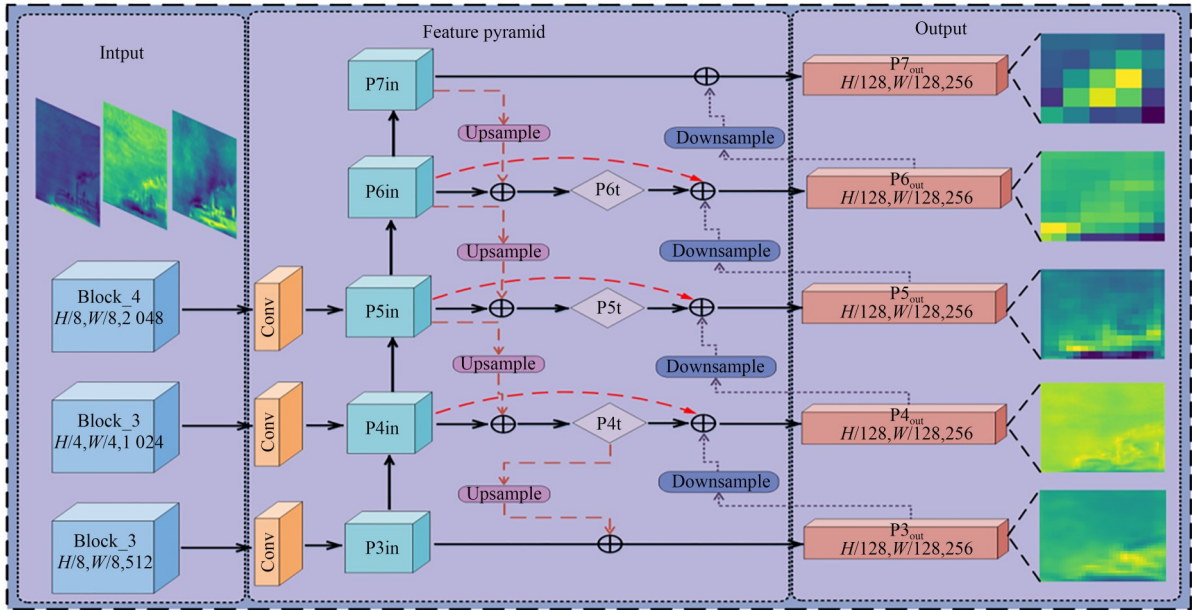


图 4 加权双向金字塔

Fig. 4 Weighted bidirectional pyramid

加权双向金字塔特征传播过程如式(2)、式(3)所示:

$$\begin{cases} P_4^1 = \text{Conv}(P_{4in}) + R(P_{5in}) \\ P_5^1 = \text{Conv}(P_{5in}) + R(P_{6in}), \\ P_6^1 = \text{Conv}(P_{6in}) + R(P_{7in}) \end{cases} \quad (2)$$

$$\begin{cases} P_{3out} = \text{Conv}(P_{3in}) + \text{Conv}(P_4^1) \\ P_{4out} = \text{Conv}(P_{3out}) + \text{Conv}(P_{4in}) + \text{Conv}(P_4^1) \\ P_{5out} = \text{Conv}(P_{4out}) + \text{Conv}(P_{5in}) + \text{Conv}(P_5^1) \\ P_{6out} = \text{Conv}(P_{5out}) + \text{Conv}(P_{6in}) + \text{Conv}(P_6^1) \\ P_{7out} = \text{Conv}(P_{6out}) + \text{Conv}(P_{7in}) \end{cases} \quad (3)$$

其中: $Pi_{in}(i=3,4,5,6,7)$ 表示输入到特征融合网络中的特征图, $P_i'(i=4,5,6)$ 表示加权双向金字塔先经过自顶向下上采样操作得到的中间特征, $Pi_{out}(i=3,4,5,6,7)$ 表示经过特征融合网络处理后的输出, $Conv(\cdot)$ 表示对特征图进行卷积处理, $R(\cdot)$ 表示对特征图进行匹配而进行的上采样和下采样操作。

为清晰解释加权双向金字塔内在工作机理对最终分割结果的有效性,在加权双向金字塔五个输出层特征图 $Pi_{out}(i=3,4,5,6,7)$ 的通道维度进行最大池化,并与原图叠加得到类激活热力

图后进行可视化对比分析。如图 5 所示(彩图见期刊电子版),图中气体区域颜色越接近红色说明模型对其激活程度越大且关注度越高。图 5 中第一行为原始特征金字塔的输出特征热力图,第二行为加权双向金字塔的输出特征热力图,从 $P3_{out}$ 和 $P4_{out}$ 热力图可以看出,加权双向金字塔输出层特征图热力分布覆盖了气体轮廓主要范围,比原始特征金字塔更能突出气体目标轮廓区域,从 $P6_{out}$ 和 $P7_{out}$ 热力图中可以看出加权双向金字塔热力分布相对于原始特征金字塔更集中在气体目标中,尤其在 $P7_{out}$ 表现更为显著。

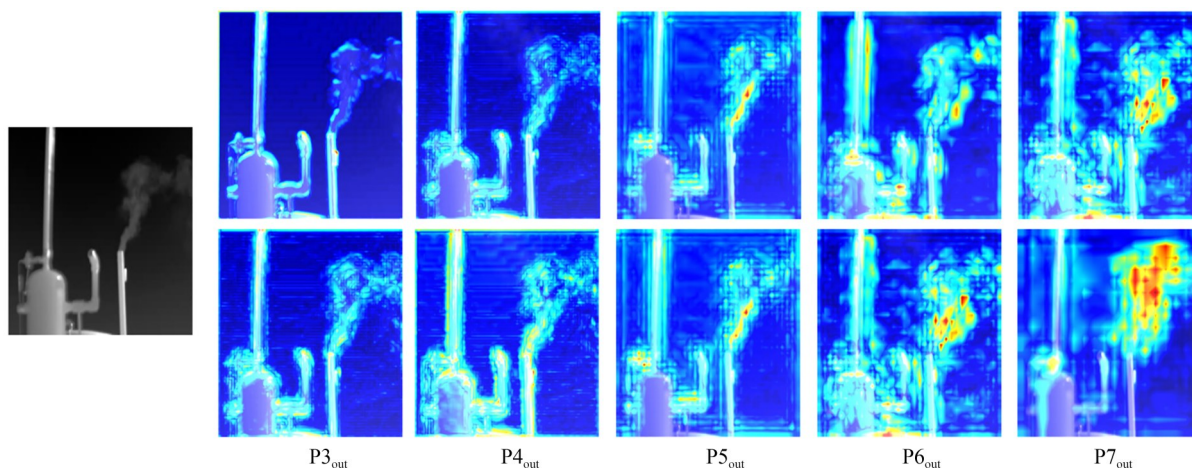


图 5 加权双向金字塔热力图对比

Fig. 5 Comparison of weighted two-way pyramidal heat map

2.4 原型特征对齐模块

Yolact 采用预测头掩码分支生成一组掩码系数,同时将金字塔 $P3_{out}$ 输出层作为与预测头并行的原型掩码分支输入,经由五层卷积组成的全卷积网络作为原型掩码分支生成通道维度为 32 的原型掩码模板,该通道维度与掩码系数个数相对应,将原型掩码模板和掩码系数以线性组合方式得到实例掩码。但 Yolact 原型掩码分支仅简单堆叠卷积层增大其感受野,忽略了不同维度特征图像之间的语义特征差异与不同尺度特征对于甲烷气体分割精度的重要性,而且甲烷气体扩散范围各异,扩散范围在图像中的尺度大小随之发生变化,这将造成甲烷气体边缘分割结果精度低、气体边界定位不准确等问题。同时考虑到单一特征表达能力的局限性,尤其是难以适应甲烷气体与背景环境低对比度和扩散区域形变复杂

的情况。本文需要在低对比度背景中对泄漏甲烷气体进行识别分割,同时甲烷气体形状、颜色、纹理特征与背景极为相似,故需要模型能捕获长距离范围内甲烷气体像素点之间的语义关系,同时也需要准确表征扩散甲烷气体的空间细节特征。为解决此问题,受文献^[29]启发本文构建了原型特征对齐模块(Prototype Feature Alignment Module, PFAM)捕获长距离甲烷气体特征之间的语义关系丰富原型语义信息量以提高生成的目标掩码质量。原型特征对齐模块如图 6 所示。

为了充分利用不同维度特征图之间的语义特征差异,使网络训练时能获得更多上下文信息,丰富各维度特征之间语义信息的关联表达,本文将加权双向金字塔 $Pi_{out}(i=3,4,5,6,7)$ 中任意两个输出特征 Pi_{out} 和 $Pi+1_{out}$ 作为原型特征对齐模块的输入特征,以捕获长距离甲烷气体特

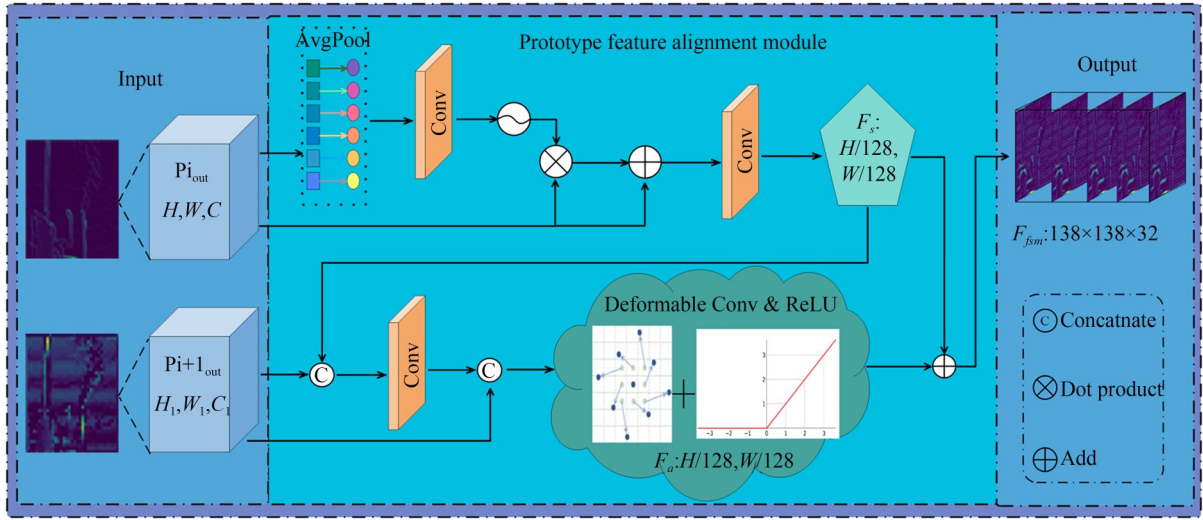


图6 原型特征对齐模块

Fig. 6 Prototype feature alignment module

征之间的语义关系、丰富各尺度特征之间语义信息的关联表达,其中 Pi_{out} 代表维度低尺度大的特征, $Pi+1_{out}$ 代表维度高尺度小的特征。原型特征对齐模块由两个分支组成,首先将 $Pi_{out}(H \times W \times C)$ 经过第一分支全局平均池化操作将特征维度从 $(H \times W \times C)$ 压缩至 $(1 \times 1 \times C)$,从而得到特征聚合后的特征向量,然后将特征向量通过卷积层实现非线性跨通道交互,并且降低其卷积计算参数量,以控制模型复杂度。最后经过 Sigmoid 函数归一化并与 $Pi_{out}(H \times W \times C)$ 点乘映射到原输入特征中,再与 $Pi_{out}(H \times W \times C)$ 进行元素相加操作实现特征二次映射,最后通过卷积层进一步融合得到特征 F_s ,实现特征选择过程,第一分支如公式(4)所示:

$$F_s = \text{Conv}_{3 \times 3} [Pi_{out} \otimes \text{Sigmoid} (\text{Conv}_{3 \times 3} (\text{Avgpool}(Pi_{out})) \oplus Pi_{out})], \quad (4)$$

其中: $\text{Conv}_{3 \times 3}$ 表示 3×3 卷积操作, Sigmoid 表示激活操作, Avgpool 表示全局平均池化操作, $\otimes$ 和 $\oplus$ 分别表示元素点乘和元素相加操作。

第二分支将特征 F_s 与 $Pi+1_{out}(H_1 \times W_1 \times C_1)$ 进行通道拼接操作聚合通道信息,经过卷积操作后与 $Pi+1_{out}(H_1 \times W_1 \times C_1)$ 再次通道拼接实现两个特征图不同通道序列间的信息交互,然后经由可变形卷积来学习特征 F_s 和 $Pi+1_{out}(H_1 \times W_1 \times C_1)$ 派生的特征偏移得到特征 F_a ,最后将具有可变形学习能力的对齐的特征 F_a

与特征 F_s 进行元素相加得到融合特征 F_{fsm} ,实现捕获长距离特征之间的语义关系达到两个不同维度特征之间特征对齐目的。第二分支如公式(5)所示,融合特征 F_{fsm} 如式(6)所示:

$$F_a = \text{DCNconv}(\text{Concat}(\text{Conv}_{3 \times 3}(\text{Concat}(Pi_{out}, F_s), Pi+1_{out}))), \quad (5)$$

$$F_{fsm} = F_a \oplus F_s, \quad (6)$$

其中: Concat 表示通道拼接操作, DCNconv 表示可变形卷积。

3 实验验证

3.1 数据集

由于目前未有公开带注释文件的红外甲烷图像数据集,本文所采用的数据集源自 GasVid^[12], GasVid 是一个由 31 个视频组成用于深度学习训练的大型甲烷视频集。数据集视频由 FLIR GF-320 红外热像仪拍摄。考虑到实验平台硬件性能瓶颈,对视频进行时域下采样稀疏处理得到 2 188 张图片作为本次的实验数据集用于训练和测试使用。本文使用图像标注工具 Labelme 根据甲烷气体泄漏区域与背景区域的像素差异性对训练集和测试集图像中甲烷气体泄漏区域进行像素级精细标注得到对应图像的真实值 (Ground Truth, GT), 保存为符合 COCO 数据集标准格式的 JSON 文件用于模型训练和测试。数据集图像及真实值标签如图 7 所示。

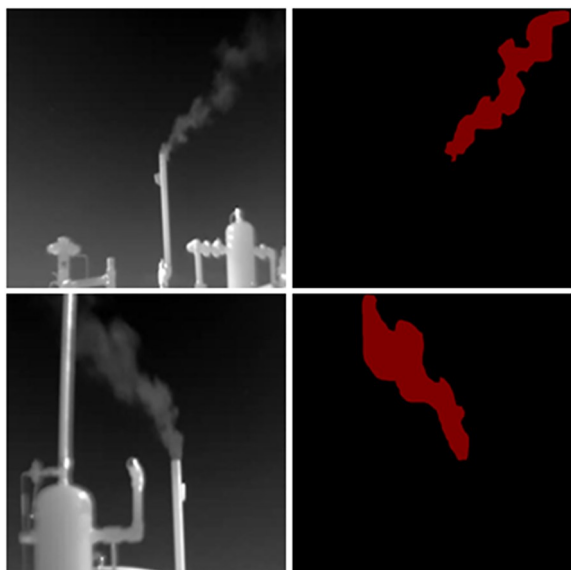


图7 数据集

Fig. 7 Datasets

实验按照 4:1 的比例将数据集随机分成训练集和测试集,分别用于训练模型和验证最终模型的泛化能力。

图 7 为数据集原图与真实值掩码,其中甲烷气体掩码边缘轮廓形变复杂,为模型掩码建模带来巨大挑战;图 8 为前景目标气体像素与背景像素占比情况,其中前景目标气体像素在图像中占比远小于背景像素占比,导致在模型特征提取过程中带来大量冗余干扰特征,加大了训练难度。

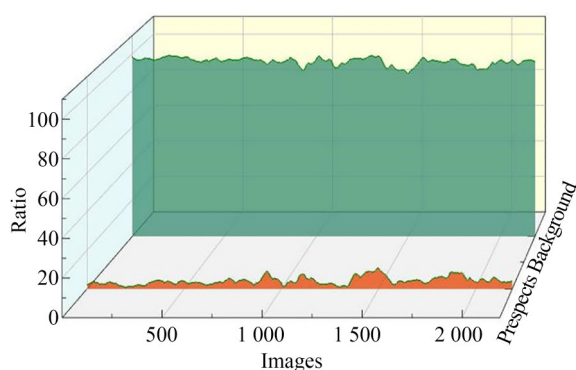


图8 前景与背景像素占比

Fig. 8 Foreground to background pixel ratio

3.2 实验环境和超参数

实验环境为 Ubuntu18.04 操作系统,Python 版本为 3.6,深度学习框架为 Pytorch 1.6.0。

CPU 为 AMD R5-3600, 16 GB 内存, GPU 为 RTX2080Ti 11 GB 显卡,使用 CUDA Toolkit 10.0 和 CUDNN V7.6.5 加速模型训练,实验各项超参数如表 1 所示。

表 1 超参数配置

Tab. 1 Hyperparameter configuration

超参数	参数值
初始学习率	0.001
学习率策略	Steps_with_decay
动量	0.9
热身迭代策略	Liner
迭代步数	120 000

3.3 评价指标

为验证本文算法优越性及有效性,本文选取平均精度(mean Average Precision, mAP)、每秒检测红外图像的张数(Frame Per Second, FPS)共两种评价指标来综合评价算法模型。考虑到实际工程应用中对甲烷气体泄漏检测有一定误报和漏报的要求,若使用 AP@95 可能会导致甲烷泄漏的漏检。综合来看,本文使用 AP50@95, AP50 来综合衡量模型性能,该评价指标能反映在不同置信度下的分割结果,是目前实例分割任务最普遍采用的评价指标。

其中 mAP 可以通过准确率(Precision)、召回率(Recall)计算得出。在 $[0.50, 0.95]$ 区间内以 0.05 为增量取 10 个交并比(Intersection over Union, IoU)作为阈值,对这 10 个阈值下对应的平均精度 AP 取平均值则得到 mAP,大小为 $[0, 1]$ 之间, mAP 越接近 1 说明模型分割效果越好。

$$IoU = \frac{P_{ij}}{\sum_{j=0}^m P_{ij} + \sum_{j=0}^m (P_{ji} - P_{ii})}, \quad (7)$$

其中: m 表示红外甲烷气体数据集中的甲烷气体类别数量, P_{ij} 为真实类别为 i 而被预测为 j 的像素数, P_{ii} 表示真实类别为 i , 被预测为 i 的像素数, P_{ji} 表示真实类别为 j , 被预测为 i 的像素数。

$$P = \frac{TP}{TP + FP}, \quad (8)$$

其中: TP (True Positive) 为真正例,即真实值是目标,预测也判别为目标, FP (False Positive) 为

假正例,真实值是目标,预测判别为不是目标。

$$R = \frac{TP}{TP + FN}, \quad (9)$$

其中:FN(False Negative)为假负例,即真实值不是目标,预测判别为不是目标。mAP 计算公式可表示为公式(10):

$$mAP = \frac{1}{m} \sum_{n=1}^m J(P, R), \quad (10)$$

其中: m 表示类别数量, n 表示类别序号, $J(P, R)$ 为准确率 P 与召回率 R 所构成 PR 曲线与横纵坐标轴所围成积分面积。

3.4 实验结果

3.4.1 自适应空间信息调控实验

自适应空间信息调控模块对主干网络中 Block_1, Block_2 和 Block_3 输出特征根据其包含空间信息量贡献程度大小差异赋予不同权重,实现低维空间信息在高级语义信息中的调控,达到对空间信息权重的自适应校准效果。为验证自适应空间信息调控模块的有效性,本组实验以 Yolact 为框架,采用 ResNet-50 和自适应空间信息调控模块作为特征提取网络进行实验分析。其中 Yolact-ASIR_1 表示自适应空间信息调控模块中采用单个自适应因子同时赋予 Block_1, Block_2 和 Block_3 三个残差块;Yolact-ASIR_3 表示自适应空间信息调控模块中采用 3 个自适应因子分别赋予 Block_1, Block_2 和 Block_3 三个残差块,实验结果如表 2 所示,各层权重对比变化如图 9 所示,后续实验在本节实验的基础上开展。

表 2 自适应空间信息调控模块实验结果

Tab. 2 Experimental results of adaptive spatial information regulation module

模型	AP50@95/%	AP50/%	FPS
Yolact	33.57	85.87	36.18
Yolact-ASIR_1	37.22	87.80	37.09
Yolact-ASIR_3	37.58	88.87	36.94

从表 2 实验结果中可得出两种结构的自适应空间信息调控模块均能对分割精度产生积极影响。采用三个自适应因子对 Block_1, Block_2 和

Block_3 特征进行空间信息调控效果相对与采用一个因子的效果好, Yolact-ASIR_3 模型 AP50 分割精度达到 88.87%, 较 Yolact 提高了 3%, 分割速度 FPS 达到 36.94。本组实验验证了自适应空间信息调控模块能有效丰富模型提取的特征信息, 自适应校准空间信息权重, 并丰富当前尺度下目标实例的特征空间和细粒度信息, 对实例目标的边缘轮廓信息量进行加强, 提高定位和分割精度。

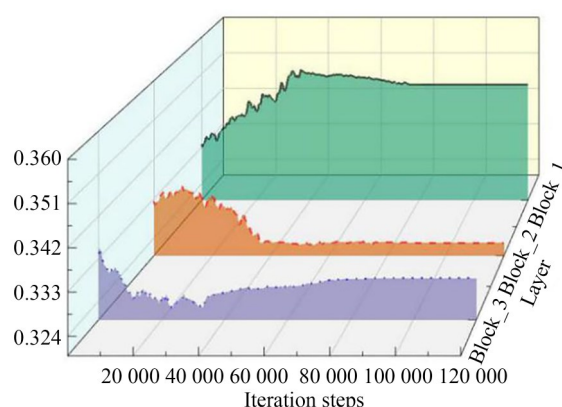


图 9 各层权重对比变化

Fig. 9 Comparison of weight changes of each layer

由图 9 可知, Block_1, Block_2 自适应因子在训练阶段变化趋势均为由大到小变化直至收敛, 在训练阶段中后期自适应因子趋于收敛状态, 而 Block_3 自适应因子在训练阶段则为小到大变化, 到训练中后期亦趋于收敛状态, 实现对 Block_1, Block_2 和 Block_3 特征信息自适应调控。

3.4.2 加权双向金字塔实验

加权双向金字塔模块通过自顶而下和自底向上两条特征传播路径实现不同维度特征融合, 减少特征横向映射过程中特征细节信息弥散丢失问题, 丰富深层特征图下气体目标实例的边缘纹理信息, 完善细粒度像素级回归任务以适应甲烷气体复杂轮廓变化下轮廓分割需求。为验证加权双向金字塔能够有效弥补特征横向映射过程中特征细节信息弥散丢失问题, 本组实验选择 Yolact-ASIR_3 模型作为基准模型, 实验结果如表 3 所示, 后续实验在本节实验的基础上开展。

表 3 加权双向金字塔实验结果

Tab. 3 Results of weighted two-way pyramid experiment

模型	AP50@ 95/%	AP50/ %	FPS
Yolact-ASIR_3	37.58	88.87	36.94
Yolact-ASIR_3-BiFPN	39.06	90.67	36.03

由表 3 可知,在引入加权双向金字塔模块后有效将深层网络的语义信息与浅层网络的空间信息进行特征融合,丰富模型提取的特征空间,带来一定计算量的提升,牺牲有限计算资源成本下使得 Yolact-ASIR_3-BiFPN 模型 AP50 分割精度达到 90.67%,在 Yolact-ASIR_3 模型基础上提升 2.02%,分割速度达到 36.03 FPS 仅下降 0.91 FPS,本组实验验证了加权双向金字塔模块对于缓解自顶而下的特征传播方式导致的低层特征空间位置和实例边缘信息弥散丢失问题的有效性。

3.4.3 原型特征对齐模块实验

原型特征对齐模块通过对不同维度特征进行特征选择和对齐学习来捕获长距离气体特征之间的语义关系,丰富原型语义信息量以提高生成目标掩码质量,实现甲烷气体分割精度提升。

为验证原型特征对齐模块能够捕获长距离甲烷气体特征之间的语义关系丰富原型语义信息量,本组实验选择 Yolact-ASIR_3-BiFPN 作为基准模型,实验结果如表 4 所示,后续实验在本节实验的基础上开展。

由表 4 可知原型特征对齐模块的输入特征对分割精度存在影响,其中将加权双向金字塔 P3out 和 P4out 输出层作为输入特征效果最好, Yolact-ASIR_3-BiFPN-FAM-P3outP4out 模型 AP50 分割精度达到 42.42%,较 Yolact-ASIR_3-BiFPN 模型升 3.36%,分割速度达到 36.80 FPS。将 P3out 和 P4out 输出层作为输入特征对分割精度影响最显著,究其原因,P3out 和 P4out 作为相邻输出层特征间语义信息关联大,训练过程中模型容易学习到两者间特征偏移关系,反之 P3out 和 P7out 特征间语义信息关联小,训练过程中难以学习到两者间特征偏移关系导致对分割精度的影响不明显。本组实验验证了特征对齐模块有效对不同维度特征进行特征选择和对齐学习来捕获长距离气体特征之间的语义关系,丰富各尺度特征之间语义信息,提高原型掩码质量实现分割精度提升。

表 4 原型特征对齐模块实验结果

Tab. 4 Experimental results of prototype feature alignment module

模型	AP50@95/%	AP50/%	FPS
Yolact-ASIR_3-BiFPN	39.06	90.67	36.03
Yolact-ASIR_3-BiFPN-FAM-P3 _{out} P4 _{out}	42.42	92.18	36.80
Yolact-ASIR_3-BiFPN-FAM-P3 _{out} P5 _{out}	40.61	91.57	37.06
Yolact-ASIR_3-BiFPN-FAM-P3 _{out} P6 _{out}	39.84	91.97	36.90
Yolact-ASIR-BiFPN-FAM-P3 _{out} P7 _{out}	38.09	90.14	37.09

3.4.4 消融实验

为了进一步验证模型中自适应空间信息调控模块、加权双向金字塔、原型特征对齐模块对于提升模型分割性能的差异,本文在自建数据集上通过消融实验验证 3 个模块组合嵌入对模型分割精度的影响,六组消融实验是单一模块或两两模块的组合算法,最后一组是本文的完整算法,实验结果如表 5 所示。实验结果表明,自适应空间信息调控模块、加权双向金字塔、原型特征对齐模块单

一或组合模块的嵌入在分割准确度上均可以产生正向优化效果,对甲烷气体分割精度都有不同程度的提升,加权双向金字塔模块的分割精度提升相对较少。自适应空间信息调控模块和原型特征对齐模块相较于 Yolact 精度提升明显,AP50@95 分割精度分别达到 37.58% 和 36.19%,在 Yolact 基础上分别提升 4.01% 和 2.62%。说明两者都可以作为模型的主要模块对特征进行处理。最后将三个模块同时嵌入模型,AP50@95 达到了

42.42%,表明三个模块同时使用对模型的提升作用更大,可以有效实现甲烷气体准确分割。

表 5 消融实验结果
Tab. 5 Results of ablation experiments

模型	ASIR	BiFPN	FAM	AP50@95/%	AP50/%	FPS
Yolact				33.57	85.87	36.18
Yolact-I	✓			37.58	88.87	36.94
Yolact-II		✓		34.94	87.71	37.38
Yolact-III			✓	36.19	88.64	39.69
Yolact-IV	✓	✓		39.06	90.67	36.03
Yolact-V	✓		✓	38.68	90.01	37.99
Yolact-VI		✓	✓	39.50	89.03	38.55
AFNet	✓	✓	✓	42.42	92.18	36.80

3.4.5 对比实验

为验证本文所提算法在自建数据集上的优越性及有效性,选择目前先进的实例分割算法 SparseInst^[30],E2EC^[31],YolactEdge^[32],SOLOv1,SOLOv2,CondInst 和 QueryInst^[33]与本文模型 AFNet 进行对比实验。为保证验证结果的有效性,实验均在 3.2 节介绍的实验平台上开展,使用的训练集和测试集均为本文自建数据集,算法性能对比如表 6 所示。从表 6 中可以看出,与其他 8 种模型相比本文设计的 AFNet 模型的 AP50@95 分割精度最高达到 42.42%,比 SparseInst、E2EC、YolactEdge 分别高出 16.97%,8.21% 和 9.63%,但分割速度低于 SparseInst 和 E2EC。对于 SOLOv1,SOLOv2,CondInst 和 QueryInst 等模型在精度和速度上均能体现出优势。综上所述,考虑到红外甲烷泄漏分割需要同时权衡模型的平均精度和推理速度,本文所设计 AFNet 模型分割性能更好,因此更适用于红外甲烷泄漏分割任务。

3.4.6 定性分析

为直观对比本文 AFNet 模型与 Yolact, SparseInst、E2EC, YolactEdge, SOLOv1, SOLOv2,CondInst 和 QueryInst 等模型分割红外甲烷图像性能的鲁棒性,将以上 9 种算法在同一测试集上对分割结果进行可视化如图 10 所示。第 1 行为原图,第 2~9 行分别为模型可视化结果图,由可视化结果得知,在(a)列图像中,YOLACT 模型将一个气体实例误检为两个气

表 6 对比实验结果

Tab. 6 Comparison of experimental results

模型	AP50@95/%	AP50/%	FPS
Yolact	33.57	85.87	36.18
SparseInst	25.45	73.17	92.88
E2EC	34.21	80.14	50.39
YolactEdge	32.79	85.30	37.35
SOLOv1	26.6	86.4	18.12
SOLOv2	30.1	85.8	12.05
CondInst	36.0	89.3	10.55
QueryInst	39.6	86.9	10.21
AFNet	42.42	92.18	36.80

体实例,而 AFNet 模型分割准确;在(c)列图像中,YOLACT 模型将背景干扰因素误检为甲烷气体。在(b)和(d)列中,对于轮廓边缘相对模糊的红外甲烷气体,AFNet 经过自适应调控主干网络空间信息,更有效的表征甲烷气体特征,因此能有效分割出甲烷气体泄漏区域,对于形变复杂的气体轮廓能做到准确拟合,排除背景和管道干扰,由此得出本文 AFNet 模型分割细节更为完善,克服背景干扰能力更强,并且误检率明显低于 YOLACT 基准模型,能够适应低对比度下甲烷气体扩散范围的尺度形状变化,整体分割结果有明显提升。

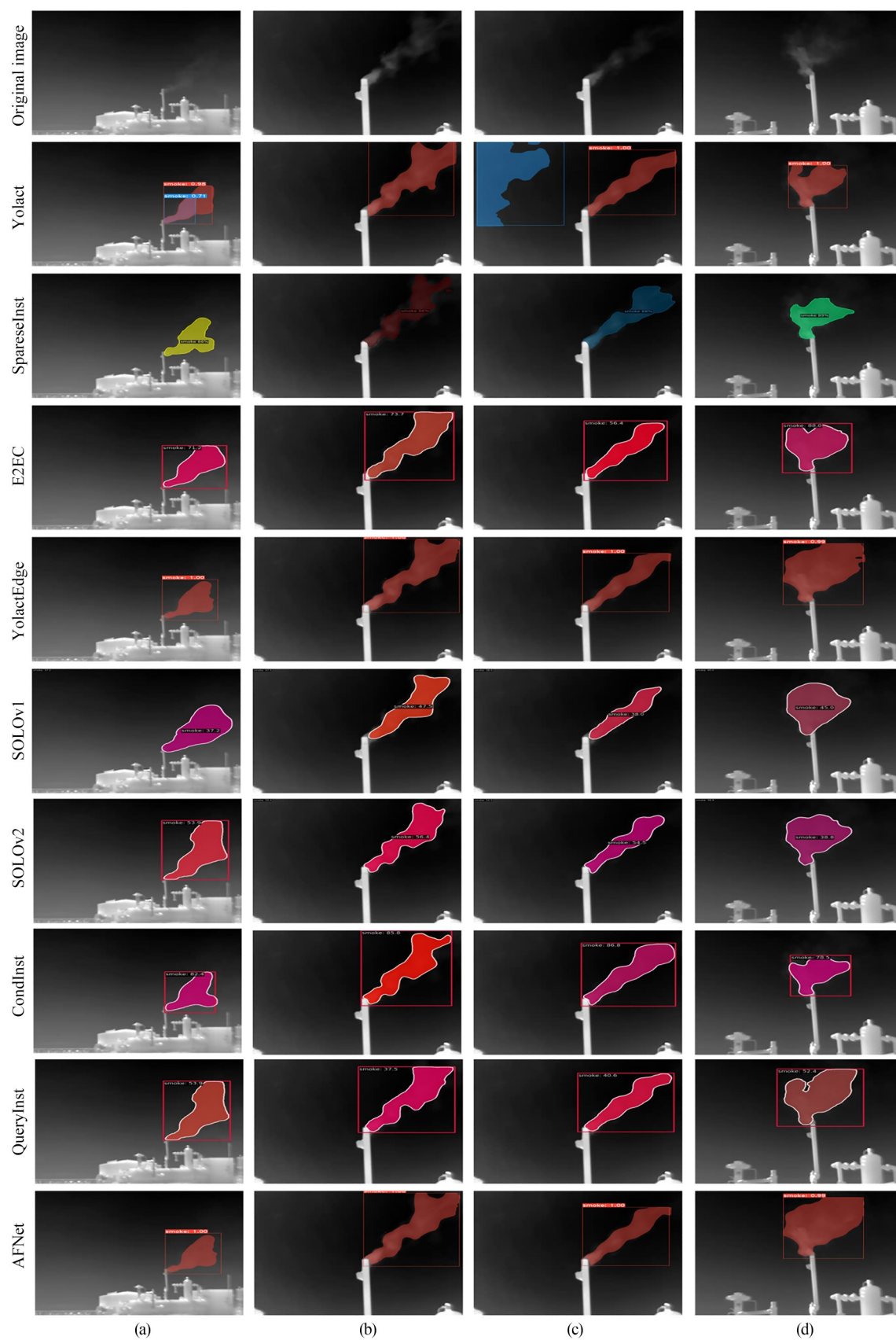


图 10 分割结果可视化

Fig. 10 Visualization of segmentation results

4 结 论

本文针对红外甲烷气体泄漏实时分割需求提出空间信息自适应调控和特征对齐模型 AF-Net,达到在低对比度背景中对尺度形变复杂的甲烷气体进行高性能分割的目的,实现在红外成像环境中对甲烷气体目标实例泄漏扩散区域准确分割。针对主干网络中高层语义特征图的空间位置信息匮乏严重的问题,AFNet模型以ResNet为主干网络,提出自适应空间信息调控模块自适应调控主干网络不同尺度残差块权重,增强骨干网络对目标特征的自适应聚合能力丰富模型提取的空间特征信息;引入加权双向金字塔弥补特征金字塔自顶而下的特征传播方式导致的目标信息逐层抽象后造成的信息弥散丢失问

题,优化模型预测层输出结果;最后构建原型特征对齐模块捕获长距离范围内甲烷气体像素点之间的语义关系,对齐不同尺度语义特征信息,丰富原型语义信息量,生成高质量掩码提高分割精度。

经过实验验证,与Yolact基础模型相比,AF-Net对甲烷气体分割精度有明显提升,掩码分割精度 AP50@95 和 AP50 分别提升 9.79% 和 6.18%,并且 AFNet 最高推理速度可达 36.80 FPS,满足化工企业在生产过程中发生甲烷泄漏事故时要求及时准确分割气体的需求,为红外甲烷气体泄漏分割提供了一定的参考价值。在未来的工作中将进一步提升精度,并且加快模型推理速度,实现在红外场景下快速和精准分割泄漏甲烷气体目标。

参考文献:

- [1] 熊仕富. 红外热成像甲烷气体探测与识别系统关键技术研究[D]. 长春: 长春理工大学, 2018.
XIONG S F. *Studies on Key Technologies of Infrared Thermal Imaging Detection and Identification System for Methane Gas*[D]. Changchun: Changchun University of Science and Technology, 2018. (in Chinese)
- [2] KEYES T, RIDGE G, KLEIN M, *et al.* An enhanced procedure for urban mobile methane leak detection[J]. *Heliyon*, 2020, 6(10): e04876.
- [3] GLOBAL METHANE INITIATIVE. Global Methane Emissions and Mitigation Opportunities [EB/OL]. Available: www.globalmethane.org, 2011.
- [4] BOAZ L, KAIJAGE S, SINDE R. An Overview of Pipeline Leak Detection and Location Systems [C]. *Proceedings of the 2nd Pan African International Conference on Science, Computing and Telecommunications (PACT 2014)*. 14-18, 2014, Arusha, Tanzania. IEEE, 2015: 133-137.
- [5] JIN H, ZHANG L, LIANG W, *et al.* Integrated leakage detection and localization model for gas pipelines based on the acoustic wave method[J]. *Journal of Loss Prevention in the Process Industries*, 2014, 27: 74-88.
- [6] YANG C B, ROMANAK K D, REEDY R C, *et al.* Soil gas dynamics monitoring at a CO₂-EOR site for leakage detection [J]. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 2017, 3(3): 351-364.
- [7] XU J H, NIE Z W, SHAN F Q, *et al.* Leak detection methods overview and summary [C]. *ICPTT 2012. Wuhan, China*. Reston, VA: American Society of Civil Engineers, 2012: 1034-1050.
- [8] SCAFUTTO RD, DE SOUZA FILHO CR. Detection of heavy hydrocarbon plumes (Ethane, propane and Butane) using airborne longwave (7.6-13.5 μm) infrared hyperspectral data [J]. *Fuel*, 2019, 242: 863-870.
- [9] HUANG K, MAO X. Detectability of infrared small targets [J]. *Infrared Physics & Technology*, 2010, 53(3): 208-217.
- [10] 丁德武, 申屠灵女, 邹兵, 等. 红外热成像技术在石化装置泄漏隐患检测中的应用[J]. *安全、健康和环境*, 2015, 15(12): 17-20.
DING D W, SHENTU L N, ZOU B, *et al.* Application of infrared thermal imaging technology in detecting leakage of petrochemical plant [J]. *Safety Health & Environment*, 2015, 15(12): 17-20. (in Chinese)
- [11] 鞠默然, 罗海波, 刘广琦, 等. 采用空间注意力机制的红外弱小目标检测网络[J]. *光学精密工程*, 2021, 29(4): 843-853.
JU M R, LUO H B, LIU G Q, *et al.* Infrared dim and small target detection network based on spatial attention mechanism [J]. *Opt. Precision*

- Eng., 2021, 29(4): 843-853. (in Chinese)
- [12] WANG J, TCHAPMI L P, RAVIKUMARA A P, *et al.* Machine Vision for Natural Gas Methane Emissions Detection Using an Infrared Camera [EB/OL]. 2019: *arXiv*: 1904.08500. <https://arxiv.org/abs/1904.08500.pdf>
- [13] XIONG K N, JIANG J B, PAN Y Y, *et al.* Deep learning approach for detection of underground natural gas micro-leakage using infrared thermal images[J]. *Sensors*, 2022, 22(14): 5322.
- [14] MELO R O, COSTA M G F, FILHO C F F C. Applying convolutional neural networks to detect natural gas leaks in wellhead images[J]. *IEEE Access*, 2020, 8: 191775-191784.
- [15] SHI J, CHANG Y, XU C, *et al.* Real-time leak detection using an infrared camera and Faster R-CNN technique[J]. *Computers & Chemical Engineering*, 2020, 135: 106780.
- [16] HE K M, GKIOXARI G, DOLLÁR P, *et al.* Mask R-CNN[C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. 22-29, 2017, Venice, Italy. IEEE, 2017: 2980-2988.
- [17] HUANG Z J, HUANG L C, GONG Y C, *et al.* Mask scoring R-CNN[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 6402-6411.
- [18] CHEN K, PANG J M, WANG J Q, *et al.* Hybrid task cascade for instance segmentation[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 4969-4978.
- [19] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. 22-29, 2017, Venice, Italy. IEEE, 2017: 2999-3007.
- [20] WANG X L, KONG T, SHEN C H, *et al.* SO-LO: Segmenting Objects by Locations[M]. Computer Vision - ECCV 2020. Cham: Springer International Publishing, 2020: 649-665.
- [21] WANG X, ZHANG R, KONG T, *et al.* Solov2: Dynamic and fast instance segmentation[J]. *Advances in Neural information processing systems*, 2020, 33: 17721-17732.
- [22] TIAN Z, SHEN C H, CHEN H. Conditional Convolutions for Instance Segmentation[M]. Computer Vision - ECCV 2020. Cham: Springer International Publishing, 2020: 282-298.
- [23] BOLYA D, ZHOU C, XIAO F Y, *et al.* YOLACT: Real-Time Instance Segmentation[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 9156-9165.
- [24] 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)*. 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [25] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 10778-10787.
- [26] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24, 2016, IEEE, 2016: 640-651.
- [27] 郝帅, 何田, 马旭, 等. 动态特征优化机制下的跨尺度红外行人检测[J]. *光学精密工程*, 2022, 30(19): 2390-2403.
- HAO S, HE T, MA X, *et al.* Cross-scale infrared pedestrian detection based on dynamic feature optimization mechanism[J]. *Opt. Precision Eng.*, 2022, 30(19): 2390-2403. (in Chinese)
- [28] WANG K X, LIEW J H, ZOU Y T, *et al.* PAN-et: Few-shot image semantic segmentation with prototype alignment[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 9196-9205.
- [29] FAN M Y, LAI S Q, HUANG J S, *et al.* Rethinking BiSeNet for real-time semantic segmentation[C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 9711-9720.
- [30] CHENG T H, WANG X G, CHEN S Y, *et al.* Sparse instance activation for real-time instance segmentation[C]. 2022 *IEEE/CVF Conference on*

- Computer Vision and Pattern Recognition (CVPR)*. 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 4423-4432.
- [31] ZHANG T, WEI S Q, JI S P. E2EC: an end-to-end contour-based method for high-quality high-speed instance segmentation [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 4433-4442.
- [32] LIU H T, RIVERA SOTO R A, XIAO F Y, *et al.* YolactEdge: real-time instance segmentation on the edge[C]. 2021 *IEEE International Conference on Robotics and Automation (ICRA)*. May 30-June 5, 2021, Xi'an, China. IEEE, 2021: 9579-9585.
- [33] FANG Y X, YANG S S, WANG X G, *et al.* Instances as queries [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. 10-17, 2021, Montreal, QC, Canada. IEEE, 2022: 6890-6899.

作者简介:



何自芬(1976—),女,山西阳泉人,博士,副教授,硕士生导师,2000年和2005年分别于西安理工大学获得学士学位和硕士学位,2013年于昆明理工大学获得博士学位,主要从事图像处理和机器视觉等方面的研究。E-mail:zyhhzf1998@163.com

通讯作者:



张印辉(1977—),男,河北故城人,博士,教授、博士生导师,2000年和2005年分别于西安理工大学获得学士学位和硕士学位,2010年于昆明理工大学获得博士学位,主要从事图像处理、机器视觉及机器智能等方面的研究。E-mail:yinhui_z@163.com