

FSNet: 基于频率特性的烟雾图像分割网络

吴健辉¹, 何 灿¹, 何俊康¹, 谢永芳², 赵 林^{1†}, 张国云¹

(1. 湖南理工学院 信息科学与工程学院, 湖南 岳阳 414006; 2. 中南大学 自动化学院, 湖南 长沙 410083)

摘要: 烟雾图像分割是对烟雾进行识别与精准定位的基础, 是火灾预警的重要手段. 针对烟雾分割时存在过分割、欠分割以及边界拟合粗糙的问题, 本文提出一种基于频率分离特性的烟雾图像分割网络. 所提出的频率分离模块将特征图中的烟雾区域分离为低频主体部分和高频边界部分, 同时基于多任务学习设计多模块权重自适应损失函数对烟雾整体、主体、边界分别监督学习, 起到细化烟雾边界和改善烟雾整体分割结果的作用; 此外, 结合可变形卷积提出改进的空洞空间金字塔池化模块以解决其信息利用率低和特征关联性差的问题. 在对比实验中, FSNet的烟雾交并比为76.55%, 比基线网络提高了4.25%. 可视化分割结果可以看出, FSNet能有效缓解过分割、欠分割, 所得烟雾边界更平滑, 烟雾图像分割的整体性能获得较大提升.

关键词: 烟雾图像分割; 频率特性; 空洞空间金字塔池化; 多任务学习; 深度学习

引用格式: 吴健辉, 何灿, 何俊康, 等. FSNet: 基于频率特性的烟雾图像分割网络. 控制理论与应用, 2023, 40(4): 702 – 712

DOI: 10.7641/CTA.2022.10857

FSNet: A network for smoke image segmentation based on frequency characteristics

WU Jian-hui¹, HE Can¹, HE Jun-kang¹, XIE Yong-fang², ZHAO Lin^{1†}, ZHANG Guo-yun¹

(1. School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang Hunan 414006, China;
2. School of Automation, Central South University, Changsha Hunan, 410083, China)

Abstract: Smoke image segmentation is the basis of smoke recognition and accurate localization, and is an important tool for fire warning. In existing algorithms for smoke image segmentation, there are some problems such as over-segmentation, under-segmentation, and rough boundary fitting. To mitigate these problems, a smoke image segmentation network based on the frequency separation characterization called as frequency separation network (FSNet) is proposed in this paper. The proposed frequency separation module divides the feature map of smoke region into a main body of low-frequency and a boundary of high-frequency. Also, a multi-module loss function with adaptive weights based on the multi-task learning is designed to learn the features of main body, boundary and whole smoke region under the supervision of the loss, refining the smoke boundary and improving the overall segmentation results of the smoke image. In addition, a modified atrous spatial pyramid pooling module is proposed to solve the problems of low information utilization and poor feature correlation by combining the deformable convolution. In the comparison experiments, the intersection over union (IoU) of FSNet is 76.55%, which is 4.25% higher than the baseline network. The visualization results show that the FSNet can effectively mitigate over-segmentation and under-segmentation, which helps to obtain smoother smoke boundary and significantly improves the overall performance of smoke image segmentation.

Key words: smoke image segmentation; frequency characteristics; atrous spatial pyramid pooling; multi-task learning; deep learning

Citation: WU Jianhui, HE Can, HE Junkang, et al. FSNet: A network for smoke image segmentation based on frequency characteristics. *Control Theory & Applications*, 2023, 40(4): 702 – 712

收稿日期: 2021–09–10; 录用日期: 2022–06–28.

†通信作者. E-mail: zhaol@hnist.edu.cn; Tel.: +86 15207302866.

本文责任编辑: 张丽清.

湖南省自然科学基金项目(2019JJ40110, 2019JJ40104), 湖南省教育厅科研基金项目(18B349, 19A201, 20A223)资助.

Supported by the Natural Science Foundation of Hunan Province (2019JJ40110, 2019JJ40104) and the Scientific Research Fund of the Hunan Provincial Education Department(18B349, 19A201, 20A223).

1 引言

在各种灾害中,火灾是发生率较高且对公众安全威胁极大的灾害之一^[1].在火灾发生的早期,物体的不完全燃烧会不断地产生烟雾,如果能在产生烟雾而没有明火之前采取灭火措施,就可以降低火灾发生的可能性,减小人民生命财产的损失.一直以来,烟雾检测是火灾预警中的重要研究内容.而烟雾分割作为更精准的烟雾检测,可对烟雾进行像素级的定位与识别,实现烟雾形态变化的动态监测,为预估火灾的扩散趋势和蔓延速度提供更为精细化的数据支撑,对火灾的分级预警以及后续的人员疏散和财产转移有着非常重要的意义^[2-3].

整体上,基于图像的烟雾检测从粗到细可分为:烟雾(有无)分类、烟雾检测和烟雾分割.从方法角度,烟雾检测方法可以概括为传统方法和基于深度学习的方法^[4].传统的烟雾检测是一种利用模式识别判断图像中烟雾有无的应用.文献[5]通过模糊C均值算法提取烟雾的颜色和纹理特征并使用支持向量机(support vector machines, SVM)分类器来分类,文献[6]采用多尺度提取烟雾的层级纹理信息,再利用SVM分类器对烟雾图像分类.但传统方法存在特征选取适用性低、计算冗余量大、时间复杂度高、算法鲁棒性低等缺点.

随着深度学习的发展,利用卷积神经网络进行分类检测任务得到了广泛关注^[7].基于深度学习的烟雾图像检测任务大体可分为3类:1)对图像进行二分类,判断图像中是否含有烟雾.如文献[8]提出深度归一化网络对烟雾图像进行分类,文献[9]提出深度多尺度网络进行烟雾识别.但当图像中烟雾比例较小或者环境较为复杂的时候,对烟雾图像进行二分类的这种粗粒度处理方式往往效果不佳;2)对图像进行检测,判断图像中是否有烟雾,有烟雾则用矩形框标定烟雾具体位置.如文献[10]提出结合更快速区域卷积神经网络(faster region-based convolutional neural network, Faster R-CNN)^[11]算法和三维卷积神经网络(three dimensional convolutional neural network, 3DCNN)^[12]的烟雾检测算法.文献[13]提出基于循环卷积神经网络的烟雾检测网络.但目标检测的检测对象多为刚性物体,对于烟雾这种非刚性的形状不规则且颜色半透明的物体检测效果不理想;3)对图像进行逐像素分类,判断图中像素点是否属于烟雾,若该像素为烟雾则区分标记.如文献[14]提出双路的深度烟雾分割网络(deep smoke segmentation, DSS)用于烟雾的精细分割,两条不同深度残差网络组成的非对称编码-解码网络提取到不同语义层次信息的特征图,最后采用加法融合策略将特征图结合并得到最后的烟雾分割结果.文献[15]提出波形网络(wave-shaped network, WNet),波谷位于波形网路编码路径的最底层,包含丰富的全局上下文信息,可以产出粗略的掩膜;波峰位于编码路径

的最顶层,包含更多细节特征,可以产出较高质量的掩膜,该网络可用于烟雾密度估计和烟雾分割.相比于烟雾有无分类和烟雾检测,烟雾分割可以减小误报率,能对烟雾更为精准的识别定位,从而实现烟雾形态变化的动态监测,具有更大的研究价值和现实意义.

烟雾是一种细小颗粒的集合,在空间中表现为不定形、透明、边界模糊的状态特点.虽然烟雾分割是一种细粒度的检测方式,可以学习到烟雾像素级的特征并根据该特征确定烟雾的具体位置,但现有的烟雾分割方法仍存在过分割、欠分割以及边界拟合粗糙等问题.在语义分割中,添加边界监督信息可以较好解决错误分割和边界拟合粗糙的问题.文献[16]将边界神经场引入全卷积网络以增强语义分割的一致性以改善目标定位效果.文献[17]合并低级特征并设计专门处理边界信息的形状流来提高边缘分割精度.文献[18]分别对目标的主体和边缘建模以取得更好的目标内部一致性和目标边界.考虑到边缘信息的重要性,以烟雾的中央主体部分低频和边界边缘部分高频的特性为基础,提出一种基于频率分离特性的烟雾图像分割网络(frequency separation network, FSNet)来改善烟雾分割质量,主要包括以下3个内容:

1)提出一种改进的空洞空间金字塔池化模块(modified-atrous spatial pyramid pooling, M-ASPP).考虑到烟雾的非刚性特征,在ASPP模块中添加可变形卷积;为提高信息利用率和特征的关联性,将ASPP模块并行处理输入特征图的方式更改为串行加并行的方式,同时通过添加 1×1 卷积和减少卷积核的数量以降低模型复杂度和减小计算开销;

2)提出频率分离(frequency separation, FS)模块.根据烟雾的频率高低将特征图分离为中央低频部分和边界高频部分,低频部分通过学习烟雾内部的语义流获得,高频部分则通过烟雾整体减去低频部分产生,此外,高频部分产生的过程中融入暗通道先验信息^[19]和底层语义信息以达到更好的效果;

3)提出多模块权重自适应损失函数监督模型训练.本文所提的烟雾图像分割算法基于多任务学习思想设计,FSNet的部分隐藏层参数共享^[20],而烟雾的低频部分和高频部分以及网络最后的输出部分通过不同的损失函数分别进行监督学习.

2 基于频率特性的烟雾图像分割网络

本文提出的烟雾图像分割网络FSNet如图1所示,总体可分为编码器、频率分离器和解码器3部分,分别对应图上、中、下3个虚线框,图中红色线条表示暗通道图像通过的路径,黑色线条表示原图通过的路径.其中,编码器由含空洞卷积的深度残差网络ResNet50和M-ASPP构成,频率分离器主体为FS模块,解码器

则主要包含相加、上采样等操作. 首先, 烟雾图像输入编码器输出具有低级语义信息的底层特征图 F_{L1} 和融合高级语义信息的特征图 F , 暗通道图像输入编码器第1层得到特征图 F_D , 然后, FS模块利用编码器的输出将烟雾分离成烟雾低频特征图 F_L 和烟雾高频特征图 F_H , 最后, 两个同形特征图 F_L 和 F_H 相加得到特征

图 F_f , F_f 与特征图 F 先拼接再上采样得到特征图 F_F 并进一步处理得到烟雾分割结果. 总体来说, 编码器提供丰富且精准的语义信息, 频率分离器则充分利用这些语义信息将烟雾分离为低频和高频特征, 解码器利用低频和高频特征以及特征图 F 得到精准的烟雾分割结果.

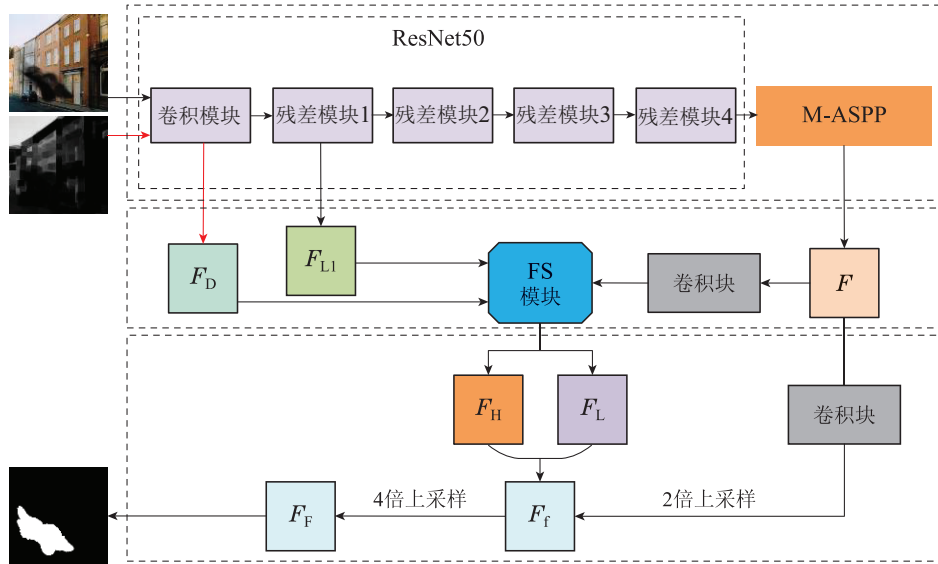


图1 FSNet网络结构示意图

Fig. 1 Schematic diagram of FSNet network structure

2.1 改进的空洞空间金字塔池化模块

本文提出的FSNet基于由ResNet50和ASPP模块构成的DeepLabV3+网络^[21]实现, ResNet50作为骨干网络提取图像的语义信息, ASPP模块则用不同采样率 r 的空洞卷积对输入特征图进行并行采样, 起到用不同感受野捕捉输入特征图的上下文语义信息的作用, ASPP模块对分割结果有着重要的意义. 本文结合烟雾的不定形特性提出M-ASPP模块, ASPP模块和M-ASPP模块如图2所示. M-ASPP模块的改进包含以下3个部分:

- 1) 针对烟雾的非刚性特性, 同时为提高模型对于

比例尺、纵横比、旋转等映射的泛化能力, 在ASPP模块中添加可变形卷积. 标准卷积中的规则格点采样导致网络难以适应几何形变. 此外, 3×3 卷积对图像进行采样提取特征时, 在同一模型的一层网络的感受野是一样的. 空洞卷积虽然通过添加空洞使得该卷积感受野变大, 但是它仍不能很好的去适应目标的空间变化^[22]. 可变形卷积通过对卷积核中每个采样点的位置增加一个可学习的偏移的变量使得该卷积核可以在当前位置附近随意的采样, 而不再局限于之前的规则格点, 进而达到不同位置的卷积核采样点位置会根据图像内容发生自适应变化的效果.

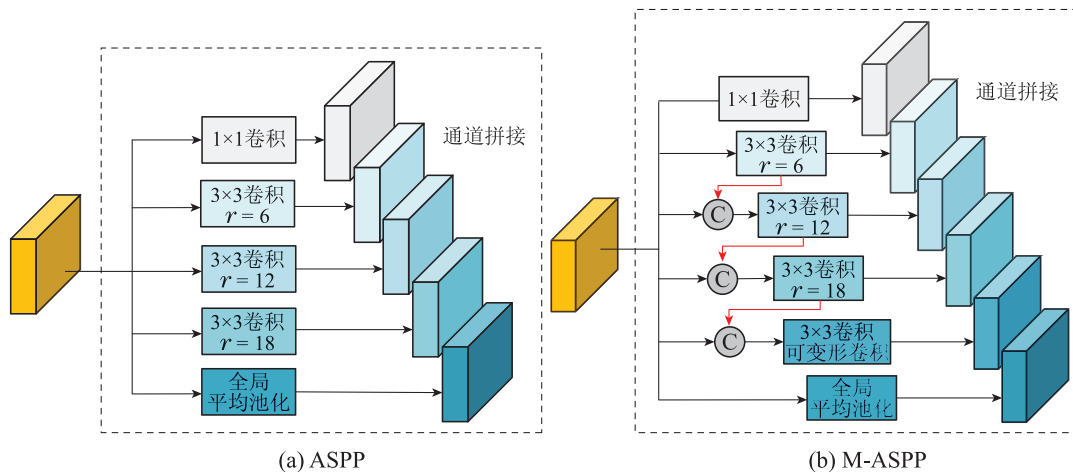


图2 ASPP与M-ASPP

Fig. 2 ASPP and M-ASPP

2) 为提高信息利用率、增加特征图的关联性, 将ASPP模块并行处理输入特征图的方式更替为串行加并行的方式. ASPP模块直接对输入特征图并行处理, 即同一特征图并行输入到不同卷积或者池化中, 这使得ASPP存在信息量丢失、信息利用率低、特征图关联性差的问题^[23], 本文将空洞卷积得到的特征图在原来结构的基础上与其他特征图拼接, 具体的, 以ASPP空洞率为6, 12的两支路卷积为例, 如图3(a)所示, 将原始特征图与 $r = 6$ 空洞卷积处理过的特征图拼接后再传入 $r = 12$ 空洞卷积, 效果如图3(b)所示, 串行加并行的方式将使得空洞卷积的感受野和有效采样点发生变化.

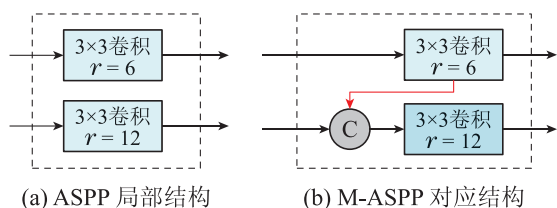


图3 ASPP与M-ASPP的局部对比

Fig. 3 Local comparison of ASPP and M-ASPP

在文献[24]中, 普通空洞卷积1维感受野 a 的计算如式(1)所示:

$$a = (r - 1) \times (k - 1) + k, \quad (1)$$

式中: r 为空洞率, k 为卷积核大小. 对于空洞卷积拼接时感受野也有相应的计算公式, 如式(2)所示:

$$a = a_1 + a_2 - 1, \quad (2)$$

式中 a_1, a_2 分别对应参与拼接的空洞卷积的感受野.

综合式(1)–(2), 空洞率 $r = 12$ 的空洞卷积1维感受野为25, 按照图3(b)拼接之后感受野为37.

空洞卷积的有效采样点是变化的, 单独的空洞卷积有效采样点为9, 而当空洞卷积进行拼接时则进行更为密集的采样^[24–25]. 图4中(a)和(b)有色点表明了二者有效采样点的分布变化.

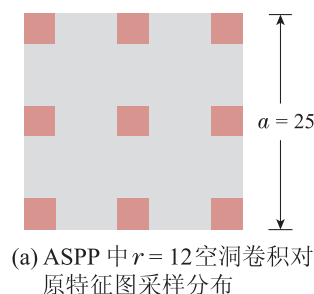
综上所述, 得M-ASPP中参与拼接的空洞卷积的信息利用率如表1所示.

从表1可知, 空洞率 $r = 12$ 和 $r = 18$ 空洞卷积拼接前信息利用率分别为1.44%和0.66%, 拼接之后分别提升2.145%和1.52%, 说明相较于ASPP, M-ASPP模块对空洞卷积的特征进行拼接可以提高信息利用率.

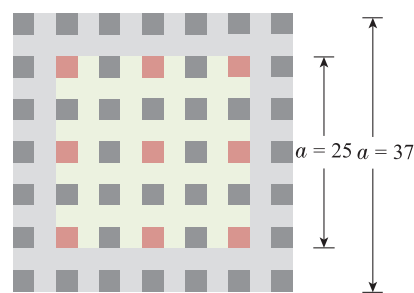
此外, ASPP中仅并行提取输入特征图的信息最后从通道维度拼接到一起得到新的特征图, 而M-ASPP在并行过程中加入了串联拼接, 使得不同感受野的空洞卷积之间有了信息传递, 因此, M-ASPP输出的特征图自身关联性比ASPP输出的特征图更强.

3) 相较于ASPP, M-ASPP添加可变形卷积并采用串行加并行的方式融合特征图, 这两种改进方式增加

了模型的计算量, 因此在M-ASPP中引入 1×1 卷积并将卷积核的数量从256降低至128以降低模型的复杂度和参数量^[26].



(a) ASPP中 $r = 12$ 空洞卷积对原特征图采样分布



(b) M-ASPP中 $r = 12$ 空洞卷积特征拼接后的采样分布

图4 空洞卷积拼接前后的感受野以及有效采样点的变化

Fig. 4 The change in receptive field and the effective sampling points after the atrous convolution concatenated

表1 拼接的空洞卷积相较原空洞卷积的变化

Table 1 The change of the concatenated atrous convolution compared to the original atrous convolution

空洞率	2维感受野	有效采样点数量	信息利用率(%)
6	13^2	9	5.33
12	25^2	9	1.44
18	37^2	9	0.66
12+6	37^2	49	3.58
18+12	61^2	81	2.18

2.2 基于频率特性的高低频分离

在传统的图像分割中, 一张图像可以被分解为低频分量和高频分量两个部分, 前者代表图像中灰度值变化缓慢的图像主要部分, 后者则对应图像中变化剧烈的边缘等其余部分. 对于烟雾图像和烟雾特征图而言, 在非烟雾部分都视为背景的前提下, 烟雾主体部分可以看作低频部分, 烟雾边界部分则为高频部分. 为说明烟雾的可分离性, 利用高斯低通滤波器对烟雾原图和标签以及烟雾原图对应的特征图分别进行高低频分离处理, 如图5所示.

基于上述分析和文献[18]的启发, 本文提出FS模块对特征图中的烟雾区域进行高低频分离, FS模块结构如图6所示. 假设分离得到的特征图满足加法原则:

$F = F_H + F_L$, 则融合暗通道信息的FS模块对特征图处理后分别可以得到烟雾的低频和高频两个部分:

$$\hat{F} = \phi(F) + \varphi(F - \phi(F)), \quad (3)$$

式中: ϕ 和 φ 分别是低频生成模块和高频生成模块进行的运算, F 是满足加法原则的原始特征图, \hat{F} 是低频生成模块和高频生成模块各自产生的特征图叠加得到的特征图.

对于烟雾低频生成模块, 如图6红色虚线框所示. 首先, 由M-ASPP得到的高维特征图 F 通过卷积得到尺寸更小的低分辨率特征图 F_{Low} . 相较于 F , F_{Low} 包含更多烟雾的高频信息, 将 F_{Low} 特征复用将有助于语义信息的利用以及指向烟雾内部的语义流的产生. 然后, 将 F_{Low} 双线性插值至特征图 F 的尺寸并与 F 拼接, 拼接的结果先后通过注意力机制模块和 3×3 卷积得到从烟雾轮廓指向烟雾几何中心的流场^[27]特征图, 最后, 流场特征图通过可微双线性采样^[28]得到

特征图 F_L .

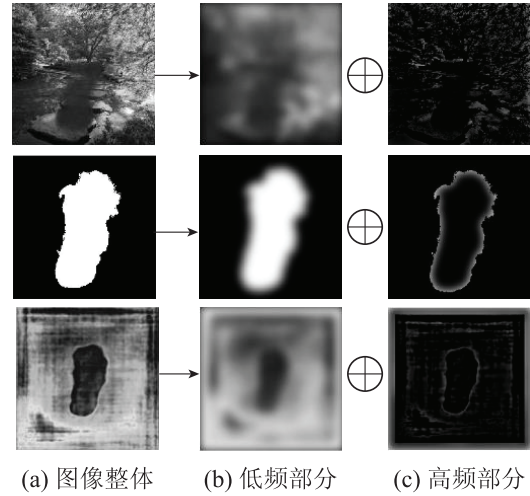


图5 传统高低频分离

Fig. 5 Traditional high and low frequency separation

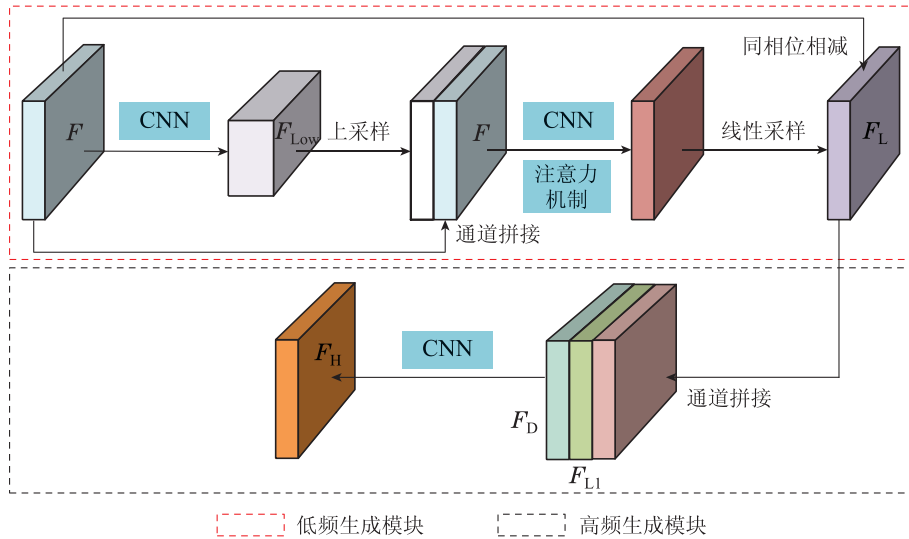


图6 FS模块结构示意图

Fig. 6 Structure diagram of FS module

对于烟雾高频生成模块, 如图6黑色虚线框所示. 首先用特征图 F 减去特征图 F_L 得到一个与烟雾高频部分相关的特征图, 然后, 将该特征图与包含细节信息的底层特征图 F_{L1} 以及由暗通道图像卷积得到的包含更多烟雾细节信息的特征图 F_D 进行拼接, 最后, 将拼接的结果通过 1×1 卷积进行降维进一步得到关于烟雾高频特征的特征图 F_H .

总的来说, FS模块对特征图 F 进行分离得到两个同形特征图 F_L 和 F_H , 二者将在解码器中进一步利用得到关于烟雾整体部分预测的特征图.

2.3 多模块权重自适应损失函数

本文所提的分割算法实质上属于多任务学习范畴, 低频生成模块、高频生成模块、以及最后的分割输出同时作用、同时受多模块权重自适应损失函数联合监

督训练. 多模块权重自适应损失函数由4部分构成, 如式(4)所示, 损失 L_L 监督低频生成模块预测烟雾低频主体部分, 损失 L_{H1} 和 L_{H2} 共同监督高频生成模块预测烟雾高频边界部分, L_F 监督网络最后的烟雾整体预测.

$$L_{Total} = L_F(F_F, G) + \omega_{H1}L_{H1}(F_H, E) + \omega_{H2}L_{H2}(F_H + F_F, G)\omega_L L_L(F_L, G), \quad (4)$$

式中: $\omega_L, \omega_{H1}, \omega_{H2}$ 为权重参数, G 指的是标注图, F_L 指的是低频生成模块所预测的烟雾主体, F_H 指的是高频生成模块所预测的烟雾边界, F_F 指的是烟雾整体预测, E 指的是预处理之后标注图的二值边界图. L_L, L_{H1} 和 L_F 均为二分类交叉熵损失函数, L_{H2} 为基于交叉熵损失的在线困难样本挖掘损失函数. 二分类交叉

熵损失函数如式(5)所示:

$$\text{Loss}(\hat{y}, y) = -\frac{1}{N} \sum_{i=1}^N (y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)), \quad (5)$$

式中: N 为图像像素总数; i 为图像上的像素点; 对于 L_L , 式(5)中 \hat{y} 与式(4)中 F_L 对应, 指预测的烟雾低频主体部分, y 与标注图 G 对应; 对于 L_{H1} , 式(5)中 \hat{y} 与式(4)中 F_H 对应, 指预测的烟雾高频边界部分, y 与 E 对应, 指的是预处理之后标注图的二值边界图; 对于 L_F , 式(5)中 \hat{y} 与式(4)中 F_F 对应, 指预测的烟雾整体, y 与标注图 G 对应。

损失函数 L_{H2} 如式(6)所示:

$$L_{H2}(F_F, G) = -\frac{1}{K} \sum_{i=1}^N G(i) \log F_F[i] \chi(i), \quad (6)$$

式中: 超参数 K 表示参与边界损失计算的像素个数, $G(i)$ 表示像素 i 对应位置的标签值, $F_F[i]$ 表示像素 i 被预测为烟雾像素的概率, $\chi(i)$ 表示值非1即0的示性函数, $\chi(i)$ 的作用是辅助挑选满足特定条件的像素 i , 具体计算公式如式(7)所示:

$$\chi(i) = \begin{cases} 1, & F_F[i] < t_K \text{ 且 } F_H[i] > t_H, \\ 0, & \text{其他,} \end{cases} \quad (7)$$

式中: $F_F[i]$ 表示像素 i 被预测为烟雾像素的概率, $F_H[i]$ 表示像素 i 被预测为烟雾边界像素的概率, 阈值 t_K 的作用是筛选 K 个预测像素作为边界像素, 阈值 t_H 的作用是筛选可靠的标签边界像素, 当网络预测的像素 i 满足示性函数的条件时才被视为网络预测的边界像素并参与损失函数 L_{H2} 的计算。

虽然多模块权重自适应损失函数由多项损失函数构成, 但是网络的最终目的为最小化输出层的损失。将 L_F 的权重系数设为1, 对于 $\omega_L, \omega_{H1}, \omega_{H2}$, 采用自适应的方式确定, 具体计算如式(8)所示:

$$\omega_j = \frac{e^{-L_j}}{e^{-L_F}}. \quad (8)$$

从式(8)可以看出, ω_j 反映该损失项对输出层损失 L_F 的影响, 当该损失对输出层是积极影响, 权值则较大; 反之, 当该损失对输出层是消极影响, 权值则较小。

FS模块将烟雾特征分离为低频部分和高频部分并通过特征复用的方式辅助预测烟雾, 而多模块权重自适应损失函数不仅监督最终的烟雾预测, 也监督烟雾高频特征和低频特征两个部分, 这种显式监督的方式使得网络更有针对性地对烟雾的边缘和主体进行学习, 得到更好的主体和边缘进而提升网络对烟雾整体的预测能力, 在一定程度上解决烟雾过度分割、分割不足以及边界拟合粗糙等问题。

3 模型分割结果与分析

3.1 实验环境设置

本文提出的烟雾分割模型基于Ubuntu18.04操作系统进行相关实验, CPU为Inter Core i9-10900x, GPU为4张NVIDIA GTX 2080Ti显卡, 训练及测试的代码均基于Pytorch深度学习框架实现。实验用Apex-Distributed Data Parallel的训练方式对模型进行训练, 训练批次设置为50, 动量为0.9, 权重衰减为0.0001, 学习率采用“Poly”策略, 初始学习率设置为0.001, 训练随着迭代次数增加而逐渐衰减学习率, 采用随机梯度下降法(stochastic gradient descent, SGD)训练模型, 每张显卡的批处理数为3张图像。

3.2 数据集

本文所用的数据集分为训练集和测试集两个部分, 训练集为5215张, 测试集为1000张^[14]。为避免过拟合, 提高模型的鲁棒性和泛化性, 本文对训练的图像和标签样本对进行预处理: 首先, 对图像进行尺寸放缩、填充裁剪和水平翻转等操作, 然后, 将数据类型转换成张量并归一化和规范化, 接着对数据进行颜色抖动, 最后, 对标签进行独热编码, 并将标签的边界放宽^[29], 求出二值边界图以配合多模块权重自适应损失函数使用。

3.3 参数分析

空洞卷积中的空洞率会影响卷积的感受野, 是影响M-ASPP效果的关键参数。将DeepLabV3+网络中的ASPP更换为M-ASPP, 其余网络结构部分与DeepLabV3+网络保持一致, 然后通过设置不同组合的空洞率来确定空洞率的组合, 实验结果如表2所示, Δ 表示其余实验组别与组别1的差值。

表2 不同组合空洞率的实验结果

Table 2 Experimental results of different combinations of atrous ratios

实验组别	不同组合的空洞率	IoU/%	Δ /%
1	$r = (1, 2, 5)$	72.04	-
2	$r = (5, 9, 13)$	72.49	0.45 ↑
3	$r = (5, 10, 15)$	72.53	0.49 ↑
4	$r = (6, 12, 18)$	72.71	0.67 ↑
5	$r = (7, 13, 19)$	72.60	0.56 ↑
6	$r = (7, 14, 21)$	72.61	0.57 ↑

从表2可以看出, 实验组别1分割结果是最差的, 这是由于空洞率直接与卷积核的感受野相关, 空洞率过小那么卷积核的感受野很小进而影响分割效果; 实验组别2、3、5、6, 分割结果略好于实验组别1, 但仍低于实验组别4的72.71。对于烟雾分割任务而言, 空洞率组合为(6, 12, 18)的卷积对于烟雾特征的提取和利用

相对恰当,因此M-ASPP中的空洞率组合设置为(6, 12, 18).

此外,对多模块权重自适应损失函数中超参数 K , t_H , t_K 进行讨论.参考文献[17–18]中对应数值的设置,先将超参数 K , t_H , t_K 分别取6500, 0.80, 0.70,然后固定其中两个超参数,调整剩余超参数记录实验结果,如图7所示.

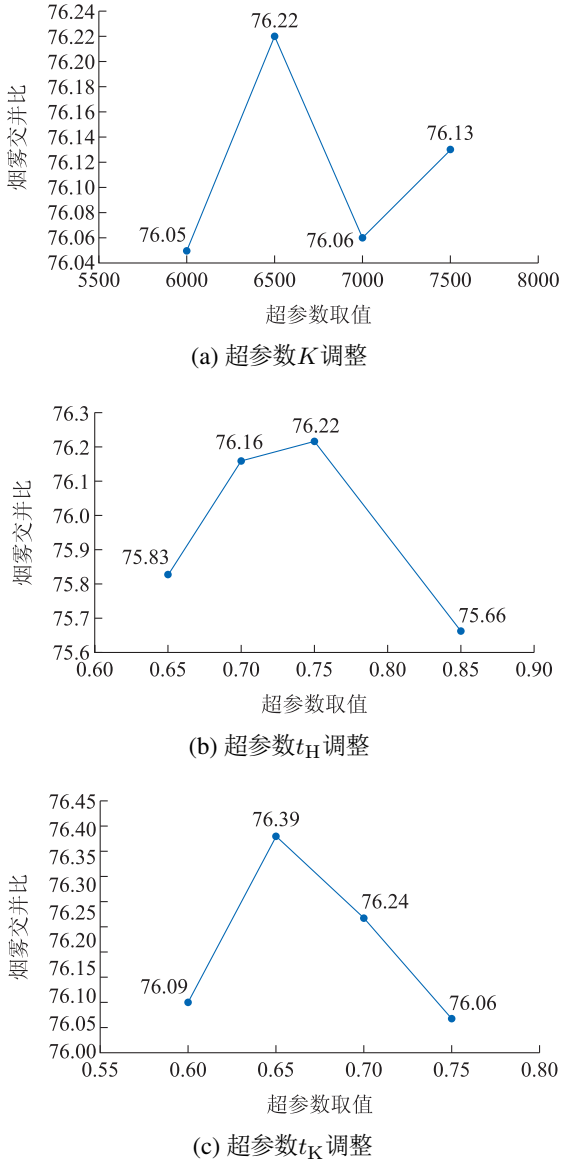


图7 超参数取值对烟雾交并比的影响

Fig. 7 Influence of hyperparameter value on smoke intersection over union

图7为单独调整超参数 K , t_H , t_K 的实验结果,从图7可以看出, K 取6500、 t_H 取0.75、 t_K 取0.65时各自模型的分割效果最好.将 K , t_H , t_K 分别设置为6500, 0.75, 0.65进行实验,测得烟雾交并比(intersection over union, IoU)为76.55%,高于图7中的76.22%, 76.39%等值,说明最优参数的组合设置对烟雾分割起到了促进作用.因此,在后续实验中超参数 K , t_H , t_K 的取值分别为6500, 0.75, 0.65.

3.4 训练可视化分析

为验证FS模块能否将烟雾特征分离为低频特征和 高频特征,对烟雾图像、对应标签以及相关特征图进行可视化,如图8所示:(a)为烟雾图像,(b)为烟雾图像的标注图,(c)为烟雾图像的主体特征图 F_L ,(d)为烟雾图像的边界特征图 F_H ,(e)为网络输出层的特征图 F_F ,也就是烟雾图像对应的整体特征图.从图8可以看出:相比(a)–(c)可以看作烟雾边界放宽后的主体特征,(d)则是烟雾边界放宽后的边界特征,说明烟雾特征是可以被分离为低频部分和高频部分的;(e)与(c)相比,烟雾特征的边界更清晰平滑,说明高低频两部分相加之后的预测结果比起单独的烟雾低频模块预测结果好.低频部分提高烟雾的内部一致性,高频部分更好的学习到烟雾的边界分布,低频部分和高频部分的相加可以使得预测结果更接近于烟雾,因此基于FS模块可以对烟雾的低频部分和高频部分进行分离并提取到有效信息,从而提升烟雾分割效果.

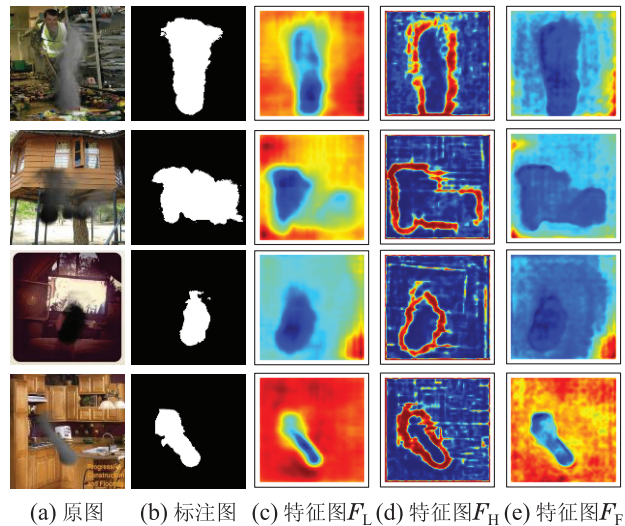


图8 可视化分析

Fig. 8 Visualization analysis

3.5 定量分析

本文用交并比和平均交并比(mean intersection over union, mIoU)作为标准来评价烟雾分割算法的性能,二者都是语义分割领域衡量算法分割效果的重要指标,交并比衡量模型对图像某个类别的分割精度,平均交并比则衡量模型对整张图像分割精度,计算公式如式(9)–(10)所示:

$$\text{IoU} = \frac{p_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}, \quad (9)$$

$$\text{mIoU} = \frac{\sum_{i=0}^k \text{IoU}_i}{k+1}, \quad (10)$$

式中: K 表示语义类别的数量,在本文中取1; i 表示像

素点的真实值; j 表示网络的预测像素值; P_{ij} 表示将类别 i 错误的预测为类别 j 的像素数量; 类似的, P_{ji} 表示将类别 j 错误的预测为类别 i 的像素数量; P_{ii} 表示将类别 i 正确的预测为类别 i 的像素数量.

搭建一些经典语义分割网络以及烟雾分割网络DSS在测试集上进行测试并求出各自的烟雾交并比, 其结果如表3所示. 相比于其他网络, FSNet在烟雾IoU数值上有较大的提升. DeepLabV3+仅取得72.30%的IoU, 而FSNet则在此基础上提升4.25%, 取得了高达76.55%的IoU. 特别的, 对于没有相应开源代码的W-Net, 先计算烟雾和背景分割的mIoU再与其原文结果进行对比, W-Net在测试集取得的mIoU为73.06%, 而FSNet在相同测试集上取得烟雾IoU 76.55%时对应的mIoU为83.89%, 相较前者高出10.83%. 综上所述, 相较于部分经典语义分割网络以及烟雾分割网络, FSNet能取得更好的烟雾分割结果.

表3 分割结果对比

Table 3 Comparison of segmentation results

方法	骨架网络	IoU/%
U-Net	CNN	55.56
DSS	VGG-16	59.65
SegNet	VGG-16	60.56
LinkNet	ResNet18	64.31
DeepLabV3+	ResNet50	72.30
FSNet	ResNet50	76.55

3.6 消融分析

为验证本文所提各个模块的有效性, 对FS模块、M-ASPP以及多模块权重自适应损失函数进行了消融分析, 以烟雾的交并比为评价指标, 实验结果如表4所示. 消融实验基于DeepLabV3+进行, 对应实验组别1. F 表示FS模块, 下面的 F_H 和 F_L 分别代表FS模块仅输出高频边界特征和仅输出低频主体特征, M 表示M-ASPP模块, $L0$ 表示二分类交叉熵损失函数, $L1$ 与 $L2$ 均为多模块权重自适应损失函数且仅有监督烟雾低频主体部分的损失函数不同, $L1$ 中监督烟雾低频主体部分的损失为边界松弛损失函数^[29], $L2$ 中监督烟雾低频主体部分的损失为二分类交叉熵损失函数, $L2$ 列下面的 L_{2-H} 代表多模块权重自适应损失函数 $L2$ 不监督低频主体部分, L_{2-L} 代表多模块权重自适应损失函数 $L2$ 不监督高频边界部分, Δ 表示其余实验组别与组别1的差值.

从表4可以看出所提模块都对分割的结果起到提升作用, 下面对消融实验进行详细分析.

1) 两两分析: 组别2相较于组别1提升0.41%, 说明M-ASPP可以更好的利用不同感受野的信息, 提取到更多烟雾特征; 组别3相较于组别1提升2.80%, 说明将

烟雾按照低频和高频的分离可以明显的提高分割的精度, 也说明分离操作可以更好的学到烟雾内部一致性和拟合出更好的烟雾边界; 组别4与组别2的区别为是否有高频边界特征参与分割, 组别4相较于组别2提升2.28%, 表明FS模块所分离出来的高频边界特征对烟雾分割是有效的; 组别5与组别2的区别为是否有低频主体特征参与分割, 而组别5相较于组别2提升3.12%, 相较于组别4的分割结果提升更多, 说明低频主体特征参与分割能大幅度提升分割效果的同时也说明烟雾的低频主体特征能比高频边界特征更多的提高烟雾分割性能, 这是因为对于烟雾而言, 低频主体部分占据烟雾整体的绝大多数比例.

表4 消融实验结果

Table 4 Results of ablation experiments

组别	F	M	$L0$	$L1$	$L2$	IoU/%	Δ /%
1	-	-	✓	-	-	72.30	-
2	-	✓	✓	-	-	72.71	0.41 ↑
3	✓	-	-	✓	-	75.10	2.80 ↑
4	F_H	✓	-	-	L_{2-H}	74.99	2.69 ↑
5	F_L	✓	-	-	L_{2-L}	75.83	3.53 ↑
6	✓	-	-	-	✓	75.66	3.36 ↑
7	✓	✓	-	✓	-	76.20	3.90 ↑
8	✓	✓	-	-	✓	76.55	4.25 ↑

2) 联合分析: 结合组别2, 4, 5, 8来看, 单独的高频边界特征或者低频主体特征对于烟雾分割都起到了较大的提升效果, 但是组别8比组别5和组别4分别高0.72%和1.56%, 意味着联合高频边界特征与低频主体特征可以更大限度的提高烟雾分割性能; 组别6相较于组别4提升0.67%, 说明相较于同时考虑烟雾以及背景两种类别的边界松弛损失函数而言, 二分类交叉熵损失更适用于烟雾分割任务; 组别8相较于组别1提升4.25%, 比组别2和组别6单独对烟雾分割提升的值加起来还多, 说明这两个模块的共同作用是正向叠加的; 组别8比起组别7进一步提升0.35%, 再次说明二分类交叉熵损失函数比边界松弛损失函数更适用于烟雾分割.

综合分析表4可知, FS模块和M-ASPP模块以及多模块权重自适应损失函数均能提升烟雾分割性能, 仅添加高频边界特征或低频主体特征也能不同程度的提升烟雾分割性能, 当高频边界特征和低频主体特征共同作用时, 烟雾分割性能提升更显著.

3.7 定性分析

为验证烟雾分割性能, 用一些经典语义分割网络、烟雾分割网络DSS以及FSNet网络在测试集上进行更为直观的定性分析, 分割结果如图9所示.

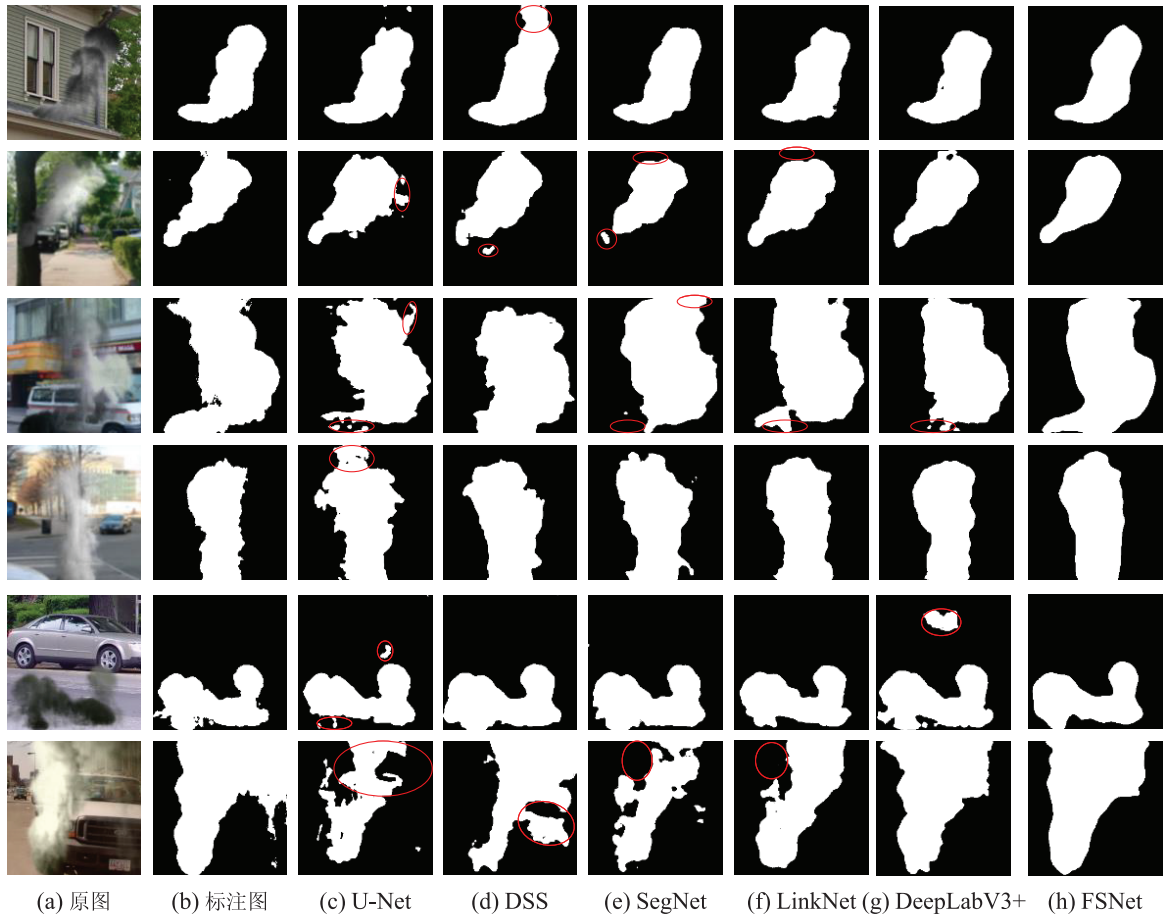


图9 不同网络分割结果对比

Fig. 9 Comparison of different network segmentation results

由图9可以看出,与标注图相比,其他的网络可能产生较大比例的错误分割或者不能很好的拟合烟雾的边界,在图9中圈出了对比网络分割的不足。U-Net网络在每幅图都有错误分割且分割结果较为粗糙;DSS网络在大部分场景分割效果要比U-Net网络要好,如第2,3,4,5,6行,但在第1行有更严重的大面积错误分割;SegNet网络和LinkNet网络相比Unet网络和DSS网络分割性能上有相对大的提升,基本正确地分割了图中的烟雾,但是也有一些过分割的问题且分割细节不够;DeepLabV3+网络整体分割性能取得了一定提升,一定程度上避免了大面积的错分割,但在第5行第7列有错误分割以及第3行第7列有较大误差的边界拟合;本文提出的网络虽然也有一定的分割不足,但相比于上述网络而言,FSNet网络极大的避免了上述网络的错误分割,分割结果在主体分割准确的前提下还取得了更精细、平滑的边界拟合,也从侧面体现了改进的空洞空间金字塔池化模块、FS模块和多模块权重自适应损失函数的有效性。

总的来说,分割结果表明,相比于其他分割网络,本文提出的FSNet网络将烟雾特征的低频和高频分离的操作,更好地保留了烟雾的边界等细节特征,同时捕捉到了更优的烟雾特征图的局部信息,即烟雾内部

一致性,可以适应图像中烟雾的形变以及区分类似于烟雾外在特征的物体,在物体高频等细节部分有着更高的语义分割准确性和更强的鲁棒性,使得烟雾分割性能得到了提升。

4 结论

针对现有的烟雾检测算法分割效果不佳以及边界拟合粗糙的问题,本文提出了基于频率特性的烟雾分割算法。对于所提M-ASPP,在ASPP模块中引入可变形卷积可以更好的应对烟雾的不定形的特性,将ASPP并行处理输入特征图的方式更替为串行加并行处理的方式使得信息利用率提高,更充分地利用不同感受野的信息。通过引入频率分离FS模块,将烟雾分离为低频部分和高频部分可以更好的学到烟雾内部一致性和拟合出更好的烟雾边界。实验结果表明本文提出的网络在多模块权重自适应损失函数监督训练下能显著提高烟雾分割的精度,并能拟合出更好的烟雾边界,为进一步的研究提供了思路,具有良好的理论和应用价值。

参考文献:

- [1] SHI Jinting, YUAN Feiniu, XIA Xue. Video smoke detection: A literature survey. *Journal of Image and Graphics*, 2018, 23(3): 303 -

322.
(史劲亭, 袁非牛, 夏雪. 视频烟雾检测研究进展. 中国图象图形学报, 2018, 23(3): 303 – 322.)
- [2] XIA Xue, YUAN Feiniu, ZHANG Lin, et al. From traditional methods to deep ones: Review of visual smoke recognition, detection, and segmentation. *Journal of Image and Graphics*, 2019, 24(10): 1627 – 1647.
(夏雪, 袁非牛, 章琳, 等. 从传统到深度: 视觉烟雾识别、检测与分割. 中国图象图形学报, 2019, 24(10): 1627 – 1647.)
- [3] CHEN Junzhou, WANG Zijie, CHEN Honghan, et al. Video dynamic smoke detection based on cascaded convolutional neural network. *Journal of University of Electronic Science and Technology of China*, 2016, 46(6): 992 – 996.
(陈俊周, 汪子杰, 陈洪瀚, 等. 基于级联卷积神经网络的视频动态烟雾检测. 电子科技大学学报, 2016, 46(6): 992 – 996.)
- [4] WANG Ziyi, SU Yuting, LIU Yanyan, et al. An improved smoke segmentation algorithm for deeplabV3 network. *Journal of Xidian University*, 2019, 46(6): 52 – 59.
(汪梓艺, 苏育挺, 刘艳艳, 等. 一种改进DeeplabV3网络的烟雾分割算法. 西安电子科技大学学报, 2019, 46(6): 52 – 59.)
- [5] ALAMGIR N, NGUYEN K, CHANDRAN V, et al. Combining multi-channel color space with local binary co-occurrence feature descriptors for accurate smoke detection from surveillance videos. *Fire Safety Journal*, 2018, 102: 1 – 10.
- [6] YUAN F N, XIA X, LI G, et al. GaborNet for smoke recognition and texture classification. *Journal of Image and Graphics*, 2019, 24(2): 269 – 281.
- [7] LI Hongxiang, WANG Xiaoli, YANG Chunhua, et al. Method of ore image segmentation based on GAN-UNet. *Control Theory & Applications*, 2021, 38(9): 1393 – 1398.
(李鸿翔, 王晓丽, 阳春华, 等. 基于GAN-UNet的矿石图像分割方法. 控制理论与应用, 2021, 38(9): 1393 – 1398.)
- [8] YIN Z J, WAN B Y, YUAN F N, et al. A deep normalization and convolutional neural network for image smoke detection. *IEEE Access*, 2017, 5: 18429 – 18438.
- [9] YUAN F N, ZHANG L, WAN B Y, et al. Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition. *Machine Vision and Applications*, 2019, 30(2): 345 – 358.
- [10] LIN G H, ZHANG Y M, XU G, et al. Smoke detection on video sequences using 3D convolutional neural networks. *Fire Technology*, 2019, 55(5): 1827 – 1847.
- [11] REN S Q, HE K M, GIRSHICK R, et al. Faster-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137 – 1149.
- [12] JI S W, XU W, YANG M, et al. 3D convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 221 – 231.
- [13] YIN M X, LANG C Y, LI Z, et al. Recurrent convolutional network for video-based smoke detection. *Multimedia Tools and Applications*, 2019, 78(1): 237 – 256.
- [14] YUAN F N, ZHANG L, XIA X, et al. Deep smoke segmentation. *Neurocomputing*, 2019, 357: 248 – 260.
- [15] YUAN F, ZHANG L, XIA X, et al. A wave-shaped deep neural network for smoke density estimation. *IEEE Transactions on Image Processing*, 2019, 29: 2301 – 2313.
- [16] BERTASIUS G, SHI J, TORRESANI L. Semantic segmentation with boundary neural fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 770 – 778.
- [17] TAKIKAWA T, ACUNA D, JAMPANI V, et al. Gated-scnn: Gated shape cnns for semantic segmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Seoul, South Korea: IEEE, 2019: 5229 – 5238.
- [18] LI X, LI X, ZHANG L, et al. Improving semantic segmentation via decoupled body and edge supervision. *Computer Vision-ECCV the 16th European Conference*. Glasgow, UK: Springer International Publishing, 2020: 435 – 452.
- [19] SHI X, LU N, CUI Z. Smoke detection based on dark channel and convolutional neural networks. *The 5th International Conference on Big Data and Information Analytics*. Qingdao, China: IEEE, 2019: 23 – 28.
- [20] ZHOU Xiaojun, GAO Yuan, LI Chao-jie, et al. End-to-end license plate recognition method based on multi-objective optimization and multi-task learning. *Control Theory & Applications*, 2021, 38(5): 676 – 688.
(周晓君, 高媛, 李超杰, 等. 基于多目标优化多任务学习的端到端车牌识别方法. 控制理论与应用, 2021, 38(5): 676 – 688.)
- [21] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European Conference on Computer Vision*. Munich, Germany: IEEE, 2018: 801 – 818.
- [22] ZHOU B, LAPEDRIZA A, KHOSLA A, et al. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(6): 1452 – 1464.
- [23] LIU Zhiyin, XIE Chunsi, LI Jinjun, et al. Smoke region segmentation and recognition algorithm based on improved Deeplabv3+. *System Engineering and Electronic Technology*, 2021, 43(2): 328 – 335.
(刘志赢, 谢春思, 李进军, 等. 基于改进Deeplabv3+的烟雾区域分割识别算法. 系统工程与电子技术, 2021, 43(2): 328 – 335.)
- [24] YANG M, YU K, ZHANG C, et al. Denseaspp for semantic segmentation in street scenes. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 3684 – 3692.
- [25] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks. *Advances in Neural Information Processing Systems*. Barcelona, Spain: NIPS, 2016.
- [26] MENG Lu, XU Lei, GUO Jiayang. A semantic segmentation algorithm based on improved mobilenetv2. *Acta Electronica Sinica*, 2020, 48(9): 1769 – 1776.
(孟磊, 徐磊, 郭嘉阳. 一种基于改进的MobileNetV2网络语义分割算法. 电子学报, 2020, 48(9): 1769 – 1776.)
- [27] DOSOVITSKIY A, FISCHER P, ILG E, et al. FlowNet: Learning optical flow with convolutional networks. *Proceedings of the IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015: 2758 – 2766.
- [28] ZHU X, XIONG Y, DAI J, et al. Deep feature flow for video recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 2349 – 2358.
- [29] ZHU Y, SAPRA K, REDA F A, et al. Improving semantic segmentation via video propagation and label relaxation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, US: IEEE, 2019: 8856 – 8865.
- [30] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Munich, Germany: IEEE, 2015: 234 – 241.
- [31] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(12): 2841 – 2495.

- [32] CHAURASIA A, CULURCIELLO E. Linknet: Exploiting encoder representations for efficient semantic segmentation. *Visual Communications and Image Processing*. Saint Petersburg: IEEE, 2017: 1 – 5.

作者简介:

吴健辉 教授, 目前研究方向为图像处理、计算机视觉、模式识别与智能系统, E-mail: jhwu@hnist.edu.cn;

何 灿 硕士研究生, 目前研究方向为图像处理, E-mail: 1023555385@qq.com;

何俊康 硕士研究生, 目前研究方向为图像处理, E-mail: 353865686@qq.com;

谢永芳 教授, 目前研究方向为复杂工业过程建模与优化控制、智能优化制造、智能控制, E-mail: yfxie@csu.edu.cn;

赵 林 副教授, 目前研究方向为智能控制、计算机视觉、复杂工业过程建模与仿真, E-mail: zhaol@hnist.edu.cn;

张国云 教授, 目前研究方向为图像处理、智能控制、计算机视觉, E-mail: gyzhang@hnist.edu.cn.