

频域多方向C-UNet及动态损失的工业烟尘图像分割

张大锦^{1,2}, 刘 辉^{1,2†}, 陈甫刚³, 赵 安^{1,2}

(1. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650500; 2. 昆明理工大学 云南省人工智能重点实验室, 云南 昆明 650500;

3. 云南昆钢电子信息科技有限公司, 云南 昆明 650302)

摘要: 工业烟尘污染等级监测中烟尘的准确分割是污染等级判定的重要前提. 针对边缘模糊且方向多变烟尘在特征提取过程中边缘方向细节信息提取困难、分割不准确的问题, 本文提出一种频域多方向C-UNet及动态损失的工业烟尘图像分割方法. 首先, 通过构建轮廓波多方向分解下采样结构增强编码阶段烟尘边缘方向信息的提取能力; 其次, 通过轮廓波变换提取烟尘8个边缘方向细节信息进行跳跃连接, 提升持续采样过程中细节信息的表达准确度; 然后, 构建轮廓波细节重构上采样结构增强解码阶段烟尘边缘细节信息的恢复能力; 最后, 提出一种动态加权策略构建组合损失函数来优化训练网络, 增强网络对烟尘边缘特征的提取能力. 结果表明, 本文方法与U-Net和其他同类方法相比在指标上有较好提升, 改善了烟尘边缘分割不准确的问题, 在不同烟尘场景上的分割效果也优于现有分割模型.

关键词: 工业烟尘; 图像分割; 轮廓波变换; 特征提取; 动态损失函数

引用格式: 张大锦, 刘辉, 陈甫刚, 等. 频域多方向C-UNet及动态损失的工业烟尘图像分割. 控制理论与应用, 2024, 41(3): 543 – 554

DOI: 10.7641/CTA.2023.20987

Industrial smoke image segmentation based on frequency domain multi-directional C-UNet and dynamic loss

ZHANG Da-jin^{1,2}, LIU Hui^{1,2†}, CHEN Fu-gang³, ZHAO An^{1,2}

(1. Faculty of Information Engineering & Automation, Kunming University of Science and Technology, Kunming Yunnan 650500, China;

2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming Yunnan 650500, China;

3. Yunnan Kunming Steel Electronic Information Technology Co., Ltd, Kunming Yunnan 650302, China)

Abstract: The accurate segmentation of the smoke in industrial smoke pollution level monitoring is an important prerequisite for pollution level determination. The typical challenges in feature extraction of smoke include blurred edges, difficult extraction of edge directional detail information and inaccurate segmentation. In this study, a frequency domain multi-directional C-UNet (Contourlet U-Net) and dynamic loss industrial smoke image segmentation method is proposed, aiming to provide support to overcome these problems. Firstly, the contourlet multi-directional decomposition down-sampling structure is constructed to enhance the ability to extract edge direction information of smoke in the encoding stage. Secondly, the contourlet transform is used to extract detailed information on the eight edge directions of smoke for skip connections, improving the accuracy of detail information expression during continuous sampling. Then, the contourlet detail reconstruction up-sampling structure is constructed to enhance the recovery ability of edge detail information of smoke in the decoding stage. Finally, a dynamic weighting strategy is proposed to construct a combined loss function to optimize the training network and enhance the network's ability to extract smoke edge features. The results show that compared with U-Net and other similar methods, the proposed method has a better improvement in indicators, improves the accuracy of smoke edge segmentation, and the segmentation effect on different smoke scenes is better than the existing segmentation model.

Key words: industrial smoke; image segmentation; contourlet transform; feature extraction; dynamic loss function

Citation: ZHANG Dajin, LIU Hui, CHEN Fugang, et al. Industrial smoke image segmentation based on frequency domain multi-directional C-UNet and dynamic loss. *Control Theory & Applications*, 2024, 41(3): 543 – 554

收稿日期: 2022-11-07; 录用日期: 2023-10-28.

†通信作者. E-mail: liuhui621@126.com; Tel.: +86 15368185621.

本文责任编辑: 周平.

国家自然科学基金项目(62263016, 61863018), 云南省科技厅应用基础研究项目(202001AT070038)资助.

Supported by the National Natural Science Foundation of China (62263016, 61863018) and the Applied Basic Research Programs of Yunnan Science and Technology Department (202001AT070038).

1 引言

工业是推进人类社会发展的基础动力,随着工业化进程的快速发展,包含大量烟尘颗粒的工业烟尘排放量逐渐增加,已成为污染环境和危害人体健康的重要来源之一.为了降低工业烟尘对人体健康的危害和环境的污染以及未来工业同生态协调可持续发展,有效控制工业烟尘的排放成为了刻不容缓的工作^[1].烟尘浓度是工业排放烟尘污染等级监测的重要评判标准,通过对烟尘排放浓度进行监控,可以有效降低大气污染程度提高环境质量.由于计算机图像识别技术对工业排放烟尘浓度进行监测具有低成本和高准确率等特点,已经成为烟尘浓度监测的主要方法^[2].其中,基于林格曼烟气黑度图的计算机图像识别方法在工业应用中使用最广,该方法包括烟尘图像采集、烟尘区域分割、烟尘区域提取、计算烟尘区域黑度值和污染等级判定5个主要环节,其中实现烟尘区域的准确分割是实现烟尘污染等级监测的难点和关键^[3].

利用计算机图像识别技术进行烟尘分割主要有传统图像分割方法和深度学习分割方法.传统图像方法主要有区域生长法^[4]、纹理分析法^[5-6]和差分法^[7].该类方法主要通过提取烟尘图像的颜色、纹理和空间结构信息等物理特征实现烟尘区域分割,存在计算冗余量大和分割精度低等问题.

随着计算机技术和性能的高速发展和快速提升,深度学习广泛应用于图像识别^[8]、目标检测^[9]和语义分割^[10]任务中,并在实际应用中取得了比传统方法更好的效果.为解决传统方法分割烟尘存在复杂度高边缘分割效果不佳的问题,Yuan等人^[11]提出一种双路径全卷积神经网络(fully convolutional network, FCN),增强了FCN对烟尘上下文和边缘特征的提取能力,但网络缺乏对烟尘形变特征的提取能力,对形状多变的烟尘边缘分割结果不够完整;因此,董泽舒等人^[12]针对形状不规则的烟尘分割困难的问题,提出了一种基于注意力机制的长距离信息建模方法,通过提取长距离像素间的依赖和连续性关系,减少了非连续区域的烟尘误判,但该方法没有针对烟尘具体特性进行分析,在小目标烟尘分割中存在漏分割现象;因而,詹光莉等人^[13]针对小目标烟尘漏分割误分割的问题使用衔接注意力机制与残差ASPP(atrous spatial pyramid pooling)的W-Net实现对小目标烟尘进行准确分割,但是该方法在提取特征的过程中方法较为单一,导致在烟尘边缘部分的分割效果欠佳;故此,吴建辉等人^[14]提出一种基于频率分离特性的烟雾图像分割网络,使用频率分离模块将特征图中的烟雾区域分离为低频和高频部分,结合频域和空间域特征细化了烟雾边界,改善了分割效果,但该方法没有考虑烟尘非刚性多方向变化特点,在烟尘边缘方向信息的提取效果上还有

待提高.以上深度学习方法在一定程度上实现了烟尘区域的分割,但都没有结合烟尘非刚性方向多变特性进行分析,缺乏对烟尘边缘方向信息的提取能力,导致对烟尘边缘区域的分割不够准确.

U-Net网络^[15]能利用全局上下文信息和在有限的训练样本下进行端到端的快速训练,能有效解决数据稀缺和类别不平衡的问题.目前,已有许多研究人员提出了基于U-Net网络的改进方案^[16-18],并取得了良好的效果.但针对烟尘图像分割任务来说,不仅需要关注烟尘分割的完整性还要关注烟尘边缘分割的准确性,对于非刚性、方向多变和边缘细节模糊的烟尘,U-Net网络还存在挖掘烟尘边缘方向信息不充分,导致烟尘边缘细节信息容易丢失分割不够准确的问题.

因此,本文针对工业烟尘非刚性多方向变化的特点,将轮廓波(contourlet)变换具有挖掘多方向信息的特性引入U-Net分割网络,构建具有提取烟尘边缘方向信息能力的C-UNet网络实现对烟尘边缘方向特征的提取和烟尘区域的准确分割.1)通过网络编码阶段的轮廓波多方向分解下采样结构,将烟尘特征分解为低频烟尘主体部分和高频多方向边缘细节分量部分,增强网络对烟尘边缘方向信息的提取能力;2)通过网络解码阶段的轮廓波细节重构上采样结构,结合跳跃连接的烟尘高频多方向细节信息和解码特征,增强网络对烟尘边缘细节信息的恢复能力;3)针对烟尘边缘特征提取困难的问题,对损失函数进行改进,提出一种基于动态加权策略的组合损失函数来优化训练网络,使网络随着训练进行从关注整体烟尘特征的提取逐渐到关注烟尘边缘特征的提取.结果表明,本文方法能够更好的提取烟尘边缘方向细节特征,对烟尘整体及边缘区域的分割准确度更高.

2 频域多方向C-UNet工业烟尘图像分割

2.1 多方向Contourlet变换

Contourlet变换是频域图像处理中一种多方向多尺度的几何分析方法^[19],能在不同尺度和不同方向上有效提取图像细节信息,将其引入深度网络中可以增强网络几何分析的能力. Contourlet变换是拉普拉斯金字塔(laplacian pyramid, LP)^[20]与方向滤波器组(directional filter banks, DFB)^[21]结合形成以轮廓段为基函数来逼近原始图像的多方向双层滤波器组结构.在不同尺度的特征图上,先使用LP分解得到一个低频子带图像和包含边缘细节的高频子带图像,再使用DFB在LP分解得到的高频子带图像上进行 l 级的多方向分解,得到 2^l 个高频方向子带.

对于一个输入图像 X ,经过LP进行多尺度分解,在第 j 层的低频子带 X_L^j 和高频子带 X_H^j 可分别表示为

$$X_L^j(m, n) =$$

$$\sum_{o=-2}^2 \sum_{p=-2}^2 M(o, p) X_L^{j-1}(2m+o, 2n+p), \quad (1)$$

$$X_H^j(m, n) = X_L^{j-1}(m, n) - X_L^j(m, n), \quad (2)$$

其中: M 为采样矩阵; m, n, o 和 p 为位置参数. 将高频子带 X_H^j 通过DFB进行方向分解, DFB分解过程中的采样矩阵可表示为如下对角形式:

$$E_k^l = \begin{cases} \text{diag}\{2^{l-1}, 2\}, & 0 \leq k < 2^{l-1}, \\ \text{diag}\{2, 2^{l-1}\}, & 2^{l-1} \leq k < 2^l, \end{cases} \quad (3)$$

其中: l 为分解等级, k 为方向角度. DFB对每一层的高频子带 X_H^j 进行分解, 将相同方向上的线性特征组合成相应的稀疏系数 s , 引入 $\phi_{k,j}$ 为DFB在不同方向上的基函数, 以内积的形式表示 s 为

$$s = \langle X_H^j, \phi_{k,j} \rangle. \quad (4)$$

图像每次经过LP分解产生的高频子带输入DFB, 逐渐将点奇异连成线型结构, 从而捕获图像中的边缘

轮廓特征. Contourlet变换具有完全重构优点, LP和DFB各自完全重建, 经对偶操作实现完全重建输入图像. 由于烟尘具有丰富的方向特征, 使用Contourlet变换进行烟尘特征的提取和重构能有效提取烟尘图像中的边缘方向细节信息, 增强烟尘边缘的分割效果.

2.2 频域多方向C-UNet

针对实际工业生产过程中, 排放的烟尘会受到多种自然环境的影响, 使烟尘图像存在边缘模糊方向多变等特点, 导致烟尘边缘分割不准确的问题. 本文通过改进原网络的编解码采样操作、跳跃连接和损失函数, 将Contourlet变换具有挖掘多方向信息的特性引入U-Net分割网络, 增强网络对烟尘图像边缘方向信息的提取能力, 提出一种频域多方向C-UNet及动态损失的工业烟尘图像分割方法, 该方法能够同时提取图像空间域和频域特征, 有效利用烟尘多方向信息且注重烟尘边缘信息提取, 整体网络结构如图1所示.

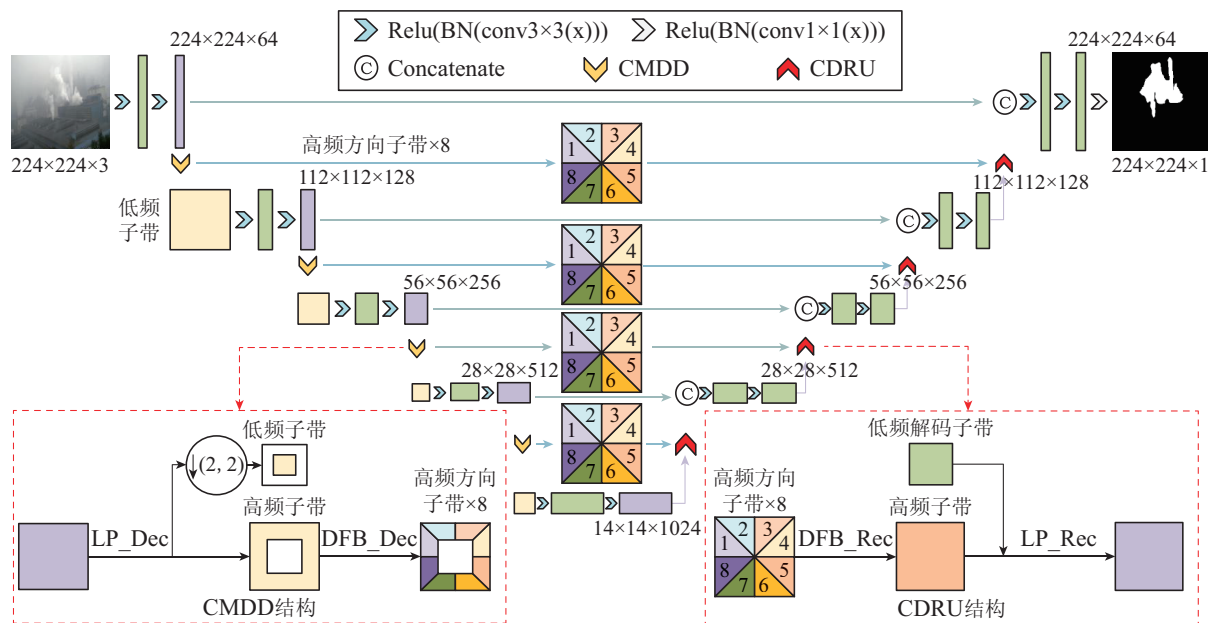


图 1 C-UNet网络结构原理图

Fig. 1 Schematic diagram of C-UNet network structure

C-UNet网络主要由编码阶段的轮廓波多方向分解下采样(contourlet multi-directional decomposition down-sampling, CMDD)、解码阶段的轮廓波细节重构上采样(contourlet detail reconstruction up-sampling, CDRU)以及多方向信息跳跃连接(multi-directional information skip connections, MISC)3个部分构成. 首先, 烟尘图像通过编码阶段实现烟尘低级和高级语义信息的提取, 其中, 在每个尺度下采样操作中, 通过使用CMDD结构得到包含烟尘图像主体信息的低频子带进行高级语义信息的提取和包含烟尘边缘细节信息的高频方向子带进行方向细节信息的传递; 其次, 在解码阶段通过结合烟尘高频多方向子带特征与低

频解码特征进行CDRU恢复特征图分辨率和烟尘边缘细节信息; 最后, 在网络中不仅对不同尺度的特征进行跳跃连接, 还对不同CMDD得到的烟尘高频多方向细节信息进行跳跃连接, 既能提高网络在持续采样过程中烟尘细节信息的表达准确性, 又能增强网络对烟尘边缘方向细节信息的恢复能力.

C-UNet网络通过在不同作用域下对烟尘图像特征进行提取, 分别对烟尘图像的低频和低频特征进行利用, 增强了网络对烟尘边缘方向细节信息的提取和恢复能力, 提高了烟尘整体及边缘区域的分割效果.

2.3 轮廓波多方向分解下采样

语义分割网络下采样操作中最大池化和平均池化

等池化操作的邻域子采样会丢失烟尘边缘细节信息,降低网络对烟尘边缘方向细节信息的处理能力.基于频域处理的轮廓波变换方法能将图像分解为包含目标主体信息的低频分量和包含目标边缘方向细节信

息的高频分量,能提高网络对细节信息的处理能力.为了增强网络对烟尘边缘方向细节信息的提取能力,本文使用轮廓波分解实现下采样操作,CMDD结构如图2所示.

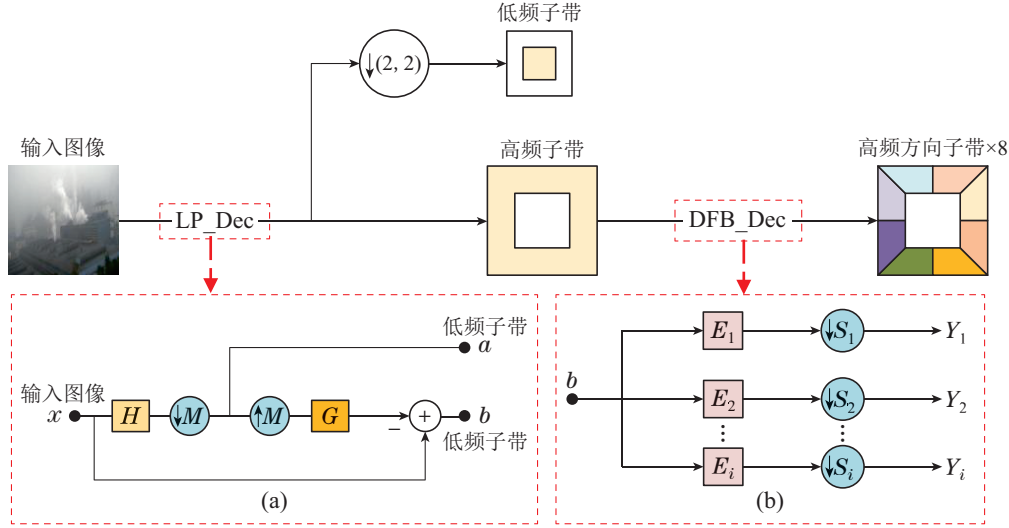


图2 CMDD结构原理图

Fig. 2 Schematic diagram of CMDD structure

在进行轮廓波分解的过程中,先通过拉普拉斯金字塔分解(LP decomposition, LP_Dec)将烟尘图像分解为一个包含烟尘主体信息的低频子带进行高级语义特征提取和包含烟尘边缘和轮廓细节的高频子带进行方向细节信息的传递,对分解得到的烟尘高频子带进行3级方向滤波器组分解(DFB decomposition, DFB_Dec)得到8个烟尘高频边缘方向子带特征.

输入图像 x 经LP_Dec之后得到低频子带 a 和高频子带 b ,即

$$a = M(H(x)), \quad (5)$$

$$b = x - G(M(a)). \quad (6)$$

图2中(a)展示了LP_Dec的详细过程,输入图像 x 经过分解滤波 H 和采样矩阵 M 下采样之后得到低频子带 a ,对低频子带 a 再经过采样矩阵 M 上采样和合成滤波 G 之后与输入图像 x 做差值处理得到高频子带 b .

高频子带 b 经过DFB_Dec分解得到8个高频方向子带 Y_i ,即

$$Y_i = S_i(E_i(b)), \quad i = 1, \dots, 8. \quad (7)$$

图2中(b)是DFB_Dec的详细过程,高频子带 b 经过分解滤波器 E 进行多方向分解后,通过采样矩阵 S 进行下采样之后得到8个高频方向子带 Y_i .

烟尘图像经过LP_Dec产生的包含烟尘边缘和轮廓的高频子带进行DFB_Dec,将LP捕捉的分布在同方向上的线性特征组合成一个轮廓段系数,形成以轮廓

段为基函数来逼近烟尘图像进行烟尘轮廓的边缘多方向信息提取.通过对输入烟尘图像进行LP_Dec和DFB_Dec之后,将输入烟尘图像的低频信息和高频信息进行分离,对烟尘高频边缘方向细节信息和烟尘低频图像主体信息分开进行特征处理,避免了在进行特征提取过程中烟尘的高频细节信息与低频主体信息互相影响,提高了网络对烟尘边缘方向细节信息的提取能力.

2.4 轮廓波细节重构上采样

实现上采样特征图分辨率恢复的常用方法主要有双线性插值、反卷积和反池化等,这些方法虽然能够恢复图像的分辨率大小,但是缺乏对图像细节信息的恢复能力.由于轮廓波变换具有完全重构的优点,能够完全重构分解图像,因此,本文在网络解码阶段通过CDRU来实现对烟尘图像特征图的分辨率恢复,CDRU结构如图3所示.

对轮廓波多方向分解下采样产生的8个烟尘高频方向子带先进行3级的方向滤波器组重构(DFB reconstruction, DFB_Rec)得到重构的烟尘高频子带特征,再结合低频解码子带特征进行拉普拉斯金字塔重构(LP reconstruction, LP_Rec),最后得到输出的重构烟尘特征.

8个高频边缘方向子带经过DFB_Rec进行完全重构得到高频子带 b ,即

$$b = \sum_{i=1}^8 D_i(S_i(Y_i)). \quad (8)$$

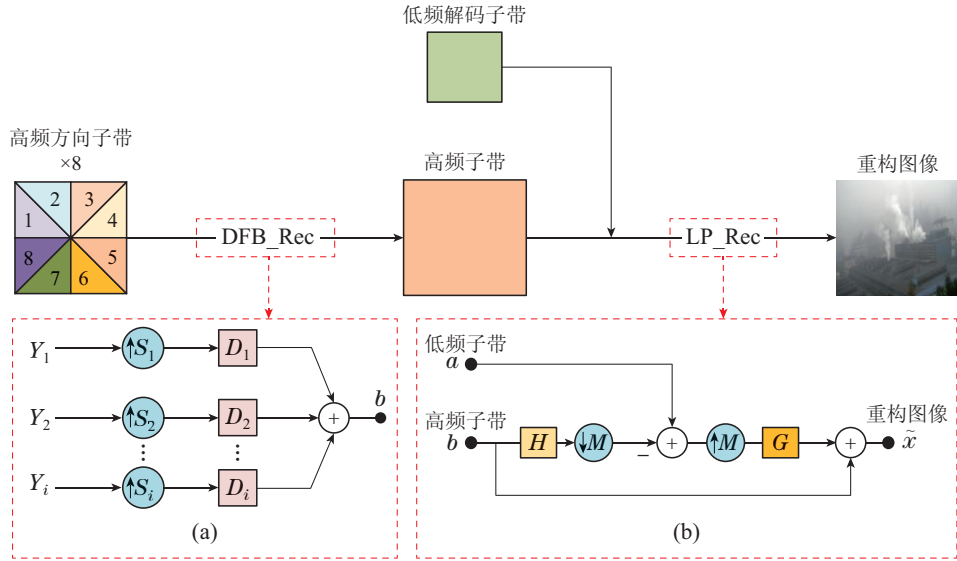


图 3 CDRU结构原理图

Fig. 3 Schematic diagram of CDRU structure

图3中(a)为DFB_Rec结构示意图, 8个烟尘方向子带先进行上采样 S 之后通过不同的方向合成滤波器 D 进行合成, 再相加融合后得到高频子带 b 。

在进行DFB_Rec之后得到的烟尘高频子带 b 再与低频解码子带 a 进行LP_Rec得到重构图像 \tilde{x} , 即

$$\tilde{x} = b + G(M(a - M(H(b)))) \quad (9)$$

图3中(b)为LP_Rec结构示意图, 高频子带 b 经过分解滤波 H 和采样矩阵 M 进行下采样之后与低频子带 a 进行差值处理得到一个近似重构烟尘图像, 再经过采样矩阵 M 上采样及合成滤波 G 之后与高频子带 b 进行相加融合, 得到最后的重构烟尘图像。

通过结合解码低频子带和在编码过程中分解出来的烟尘高频方向子带信息进行CDRU有效利用了烟尘高频方向细节信息, 实现了烟尘图像特征图分辨率和边缘细节信息的恢复, 增强了在解码过程中烟尘边缘细节信息的恢复能力。

2.5 多方向信息跳跃连接

在网络中进行编解码特征提取过程中会存在烟尘边缘细节信息丢失的情况, 为增强网络对烟尘边缘方向细节信息的提取恢复能力, 本文在网络中通过MISC实现了对烟尘边缘多方向细节信息进行单独利用. 在网络编码操作的过程中, 通过CMDD得到的烟尘高频方向子带信息, 以跳跃连接的方式直接传递到CDRU操作中, 降低了由于网络持续采样提取特征过程中造成的烟尘细节信息损失. 通过烟尘高频方向细节信息的直接传输, 增强了网络在解码过程中对烟尘图像边缘细节信息的恢复能力. 另外, 在进行解码操作的过程中, 结合高频跳跃连接的烟尘多方向细节信息和烟尘低频解码特征进行CDRU恢复烟尘图像特征, 能有效增强网络恢复烟尘边缘细节信息的能力。

C-UNet通过把包含烟尘边缘方向细节信息的高频分量和包含烟尘主体信息的低频分量分开进行处理, 使烟尘主体低频分量进行持续提取高级烟尘语义信息的过程中不会受烟尘高频边缘细节信息的影响, 同时烟尘高频方向细节信息为上采样恢复分辨率的过程中提供了烟尘多方向细节信息, 增强了网络恢复烟尘边缘方向细节信息的能力。

3 动态加权组合损失函数

在烟尘分割的过程中, 交叉熵损失(cross entropy loss, CEL)函数^[22]在网络训练的时候能够同时关注背景与待分割目标前景区域, 为所有像素保持平滑的梯度, 有助于对所有像素进行分割, CEL计算公式如下:

$$CEL = - \sum_{i=1}^m y \log \hat{y}, \quad (10)$$

其中: y 表示烟尘真实标签, \hat{y} 表示烟尘预测结果, m 表示类别总数。

但是CEL在烟尘的边缘精细部分的分割效果较差, 特别是在烟尘的边缘像素的区分上置信度较低, 导致烟尘图像边缘细节分割模糊. 针对此问题, 本文提出了一种基于动态加权策略的组合损失函数来增强网络对于烟尘边缘特征的提取。

基于结构相似性指标构建的结构相似性损失(structural similarity loss, SSIML)函数^[23]能够捕捉图像边缘结构信息, 考虑到每一个像素结构的局部领域, 在进行损失计算的时候为边界分配更高的权重, 以此在训练的时候使网络更加关注边界的精细分割, 其计算公式如下:

$$SSIML = 1 - \frac{(2\mu_y \mu_{\hat{y}} + C_1)(2\sigma_{y\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)}, \quad (11)$$

其中: $\mu_y, \mu_{\hat{y}}$ 和 $\sigma_y, \sigma_{\hat{y}}$ 为标签 y 和预测值 \hat{y} 的均值和方差; $\sigma_{y\hat{y}}$ 是标签和预测值的协方差; $C_1 = 0.001^2, C_2 = 0.003^2$.

焦点损失(focal loss, FL)函数^[24]能够通过参数调整正负样本和难易分割样本对损失函数的贡献,能在网络训练的时候使网络更偏向于难分割样本,增强网络对难分边界的分割效果,其计算公式如下:

$$\text{FL} = -\alpha(1 - \hat{y})^\gamma y \log \hat{y} - (1 - \alpha)(\hat{y})^\gamma (1 - y) \log(1 - \hat{y}), \quad (12)$$

其中: α 是样本数量平衡因子,用来调节正负样本的比重; γ 是难易平衡因子,用来调节难分样本对损失函数的贡献. 本文中设置 $\alpha = 0.75, \gamma = 2$.

普通组合损失函数通过使用固定参数对不同损失函数进行加权组合,加权参数的固定设置,不利于最佳参数的寻找和网络的训练过程. 为了使损失函数适应网络的训练过程,更能准确分割烟尘边缘及整体区域,受文献[25]的启发,对CEL函数进行改进,采用动态变化的权重参数来实现组合损失的加权,使组合损失具有自适应性和动态性,能使网络模型从关注全局信息逐渐关注烟尘边缘区域,提高烟尘边缘区域的分割效果.

本文采用lg函数构造了一种动态参数变化方案,通过设置最小权重参数和最大权重参数,使加权参数在网络训练的过程中根据迭代次数的变化进行动态调整. 以迭代次数动态变化参数 $\frac{n}{N} \in (0, 1]$ 为自变量的lg函数为

$$H = \lg \frac{n}{N} \in (-\infty, 0], \quad (13)$$

其中: n 为当前迭代次数, N 为总迭代次数.

当网络刚开始训练,即 $n = 1$ 时,需要通过本文构造的动态加权策略使组合损失函数的加权参数接近于最小权重参数;随着网络训练到最后一轮,即 $n = N$ 时,需要加权参数等于最大权重参数. 因此,要使动态参数能在最小权重参数和最大权重参数范围内动态变化,需改进自变量为 $1 + 9\frac{n}{N} \in (1, 10]$, 以得到动态变化函数为

$$H = \lg(1 + 9\frac{n}{N}) \in (0, 1]. \quad (14)$$

最后,基于上述得到的动态变化函数,即可得到本文提出的在最小权重参数和最大权重参数范围内随着网络训练的进行动态变化的动态加权策略

$$\beta_{\text{cur}} = \beta_{\text{min}} + (\beta_{\text{max}} - \beta_{\text{min}}) \lg(1 + 9\frac{n}{N}), \quad (15)$$

其中: β_{cur} 为当前迭代次数下的 β 值, β_{max} 为设定的参数最大值, β_{min} 为设定的参数最小值.

通过上述动态参数构建的动态加权组合损失函数

如下:

$$L_{\text{combination}} = (1 - \beta) \text{CEL} + \beta (\text{SSIML} + \text{FL}). \quad (16)$$

随着网络训练的进行, β 值随着动态增长变化,使动态加权组合损失函数更适应网络的训练过程,让网络从训练开始先通过CEL来关注烟尘全局信息,提取烟尘整体区域;在此基础上再通过SSIML和FL来关注烟尘难分割的边缘区域,提取精细烟尘边缘,以达到网络对烟尘整体和精细边缘区域的准确分割.

4 实验与分析

4.1 实验平台及数据

实验所用电脑配置为Intel(R)Xeon(R) CPU E5-2620 V4 @2.10 GHz处理器, 64 GB内存, 11 GB RAM的NVIDIA GeForce RTX 2080Ti显卡,实验使用Python语言, PyTorch深度学习框架.

实验所用工业烟尘图像来自实际钢厂定点拍摄的烟尘排放视频,视频记录了白天时间段内不同监测点不同场景的烟尘排放信息. 从拍摄视频中选取不同时间段、不同拍摄点的60段烟尘视频,从中提取1724帧图像,设置图像尺寸大小为 224×224 , 使用Labelme图像标注软件在专家指导下手动对烟尘目标区域进行标注,制作烟尘图像数据标签. 对烟尘图像和标签进行上下翻转和 $90^\circ, 180^\circ, 270^\circ$ 旋转实现数据集扩充至8620张,其中训练集6370张、测试集2250张. 为了验证本文方法在不同烟尘场景下的识别能力,将测试集中的2250张图像划分为昏暗烟尘场景840张、易辨烟尘场景340张、干扰场景210张、小目标烟尘场景380张、稀薄烟尘场景175张、有颜色烟尘场景305张,共6个场景.

4.2 实验设置

本文实验在PyTorch深度学习环境下,通过端到端的方式进行训练,训练使用Adam算法对损失函数进行优化,优化器参数 $\beta_1 = 0.9, \beta_2 = 0.999$, 批处理大小 $\text{batch} = 2$, 学习率 $\text{lr} = 0.0001$, 训练迭代次数 $\text{epoch} = 50$.

4.3 评价指标

为了对本文方法进行有效评估,选择了查全率 $R(\text{recall})$ 、查准率 $P(\text{precision})$ 、IOU(intersection over union)、DSC(dice similarity coefficient)和F1度量(F1-score)作为评价指标. 其中: R 是被正确分类的正样本数量与真正正样本的数量之比; P 是被正确分类的正样本数量与被预测为正样本的数量之比; IOU是语义分割的标准度量,是真实值与真实预测结果的交集与并集之比; DSC用于度量两个集合的相似程度; F1度量是综合度量查全率和查准率的指标,本文中正样本为烟尘目标区域、负样本为背景区域. 各评价指标计

算公式如下:

$$R = \frac{TP}{TP + FN}, \quad (17)$$

$$P = \frac{TP}{TP + FP}, \quad (18)$$

$$IOU = \frac{TP}{TP + FP + FN}, \quad (19)$$

$$DSC = \frac{2TP}{2TP + FP + FN}, \quad (20)$$

$$F1 = \frac{(1 + \beta^2) R \times P}{\beta^2 P + R}, \quad (21)$$

其中: TP, TN, FP, FN表示真实值与预测值之间的关系, 如表1所示. 式(21)中, β 用于调整查全率和查准率的权重, 当 $\beta = 1$ 时, 二者重要度一样; β 值增大, 则认为查全率更重要; β 值减小, 则认为查准率更重要. 在本实验中, 更关注被分割出来的烟尘正样本数量, 即认为查全率更重要, 设置 $\beta = 2$.

表 1 TP, FP, FN, TN含义表
Table 1 Meaning of TP, FP, FN, TN

混淆矩阵		预测值	
		Positive	Negative
真实值	Positive	TP	FN
	Negative	FP	TN

4.4 消融实验分析

为了验证本文所提结构上改进的有效性, 本节对U-Net^[15]、单独使用CMDD结构(Model1)、单独使用CDRU结构(Model2)、去除DFB分解和重构只进行LP分解和重构(LP-Net)和本文方法C-UNet进行对比. 在进行Model1和Model2的消融实验过程中, 网络的上采样和下采样过程使用原始U-Net网络的上采样与下采样结构. 其中, 在验证CDRU结构的有效性时, 需要使用到CMDD结构分解得到的8个高频方向子带, 因此, 在网络的下采样过程中, 仅使用CMDD结构为CDRU结构提供8个高频方向子带, 并不参与网络的训练. 各个网络训练所用损失函数为CEL, 评价指标对比如表2所示, 分割结果对比如图4所示.

表 2 网络结构评价指标对比

Table 2 Comparison of network structure evaluation indexes

模型	R	P	IOU	DSC	F1
U-Net ^[15]	0.8450	0.8617	0.7462	0.8339	0.8483
Model1	0.9080	0.8514	0.7831	0.8662	0.8961
Model2	0.9084	0.8471	0.7809	0.8632	0.8954
LP-Net	0.9018	0.8525	0.7815	0.8625	0.8915
C-UNet	0.9212	0.8583	0.7984	0.8751	0.9079

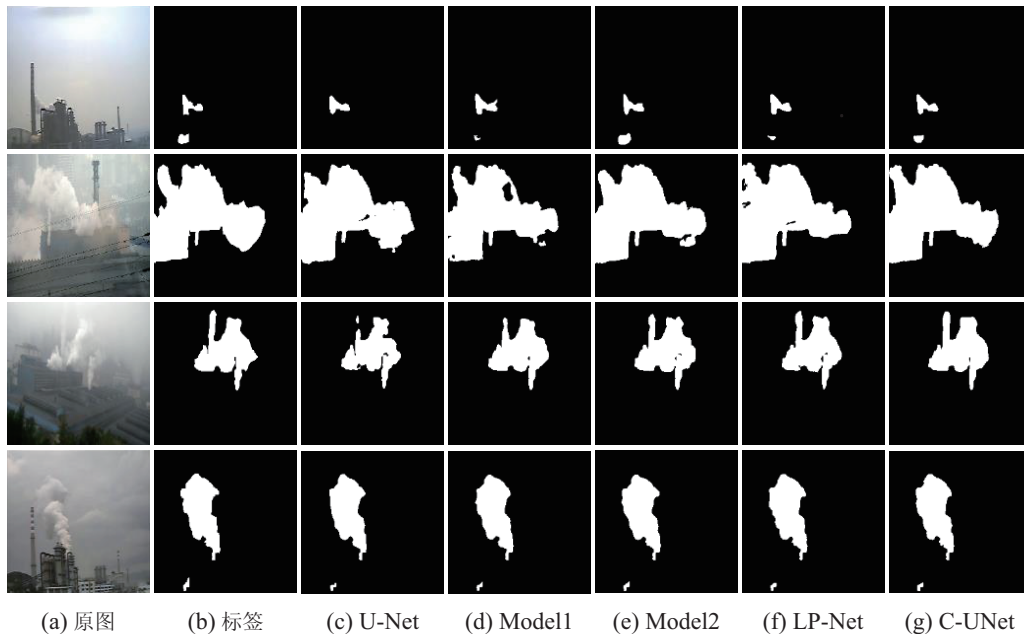


图 4 网络结构改进前后对比

Fig. 4 Comparison of network structure before and after improvement

1) 从实验结果中可以看出, U-Net中由于对烟尘整体特征和边缘方向细节信息的提取能力不足, 存在烟尘边缘细节信息丢失和烟尘漏分割的问题, 如图4中第3列的第1个图所示;

2) 单独使用轮廓波进行下采样会降低由于池化

操作的邻域子采样造成烟尘边缘细节信息的问题, 能增强网络对烟尘边缘方向细节信息的提取能力, 从表2中的评价指标也可直观看出, Model1比U-Net在R, IOU, DSC和F1这4个评价指标上分别提高了6.30%, 3.69%, 3.23%和4.78%, 但Model1在提取烟尘

整体和边缘细节信息上的能力还有待提高,对烟尘分割还不够准确,如图4中第4列的第2个图所示;

3) 单独使用轮廓波进行上采样能够利用轮廓波变换具有完全重构的优点,能够完全重构分解图像,增强网络对烟尘图像边缘细节信息的恢复能力,从表2的评价指标可以看出,Model2比U-Net在R, IOU, DSC和F1这4个评价指标上分别提高了6.34%, 3.47%, 2.93%和4.71%,但Model2同样存在分割出来的烟尘整体和边缘不够精确的问题,如图4中第5列的第4个图所示;

4) LP-Net去除了DFB分解和重构只进行LP分解和重构,在一定程度上增强烟尘边缘细节信息的提取能力,从表2的评价指标可以看出,LP-Net比U-Net在R, IOU, DSC和F1这4个评价指标上分别提高了5.68%, 3.53%, 2.86%和4.32%,但LP-Net缺乏对烟尘多方向信息的提取能力,对烟尘边缘部分的提取还有待提高,如图4中第6列的第3个图所示;

5) 本文方法C-UNet通过CMDD将图像分解为包含烟尘主体信息的低频分量和包含烟尘边缘方向细节信息的8个高频分量,有效提高了网络对烟尘细节信息的处理能力.通过CDRU结合跳跃连接的多方向细节信息,能有效提高网络对烟尘边缘方向细节信息的恢复能力,提高烟尘边缘细节部分的分割效果,如图4中第7列第4个图所示.从表2中的评价指标中可以更直观的看出,增加方向信息之后,本文方法C-UNet通过提取8个烟尘方向进行边缘细节特征提取,在R,

IOU, DSC和F1这4个评价指标上比U-Net分别提高了7.62%, 5.22%, 4.12%和5.96%,说明增加烟尘方向信息提取能有效增强网络对烟尘边缘信息的提取能力,提高网络对烟尘整体和边缘区域的分割能力.

4.5 损失函数改进对比

为了验证本文提出的动态加权损失函数的有效性,将所提的动态加权损失和其他损失函数进行对比.改进前后对比的损失函数有CEL^[23], SSIML^[24], FL^[25]以及固定参数分别为($\beta = 0.6$, $\beta = 0.7$, $\beta = 0.8$)下的组合损失函数和本文所提动态加权策略下(参数动态变化范围分别为: (0.2, 0.6], (0.3, 0.7], (0.4, 0.8])的动态组合损失函数.9个损失函数在C-UNet上的评价指标对比如表3所示,分割结果对比如图5所示.

表3 损失函数评价指标对比

Table 3 Comparison of loss function evaluation indexes

模型	R	P	IOU	DSC	F1
CEL ^[23]	0.9212	0.8583	0.7984	0.8751	0.9079
SSIML ^[24]	0.0842	0.0249	0.0143	0.0279	0.0570
FL ^[25]	0.9288	0.8004	0.7528	0.8480	0.8999
$\beta = 0.6$	0.9248	0.8507	0.7948	0.8735	0.9090
$\beta = 0.7$	0.9288	0.8517	0.7976	0.8755	0.9123
$\beta = 0.8$	0.9260	0.8501	0.7950	0.8741	0.9097
(0.2,0.6]	0.9330	0.8438	0.7937	0.8723	0.9137
(0.3,0.7]	0.9375	0.8451	0.7989	0.8770	0.9174
(0.4,0.8]	0.9316	0.8521	0.7996	0.8778	0.9146

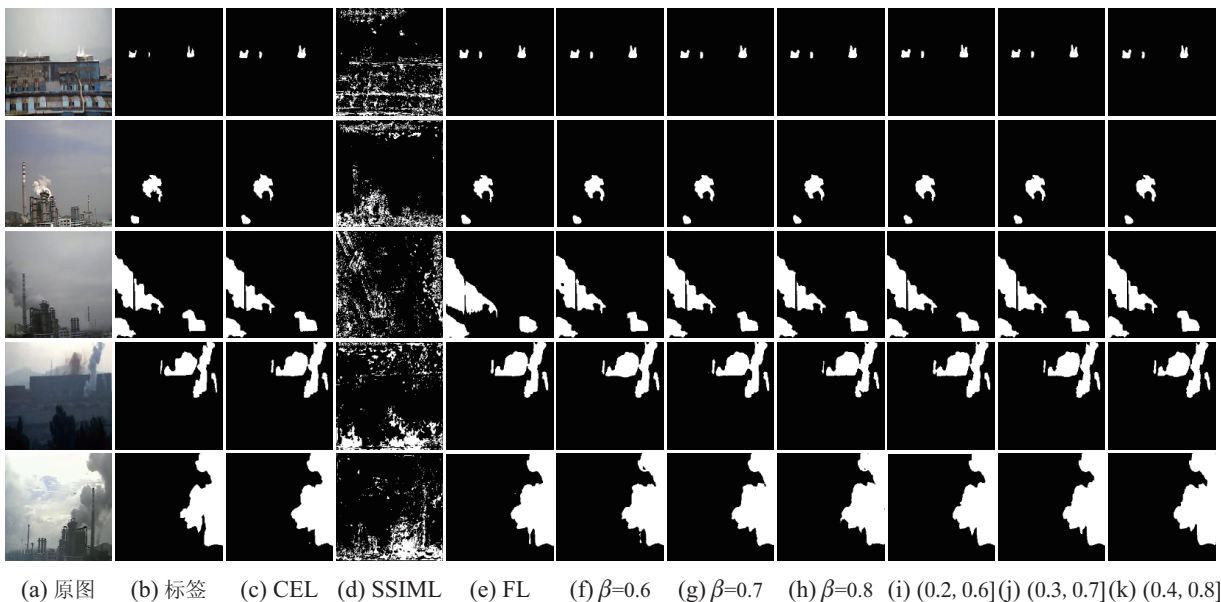


图5 损失函数实验结果对比

Fig. 5 Comparison of experimental results of loss function

1) 从实验结果可知,采用单一损失函数对网络进行训练,由于对烟尘的整体特征和边缘特征没有针对性,使网络对烟尘边缘的关注度不够,分割结果不够

精确.其中,SSIML的分割效果差的原因是单独使用SSIML的时候网络关注的是全局图像的边缘结构信息,缺乏烟尘先验信息指导,导致单一SSIML下烟尘

的分割效果较差. 从图5中(c)–(e)列可以看出单一损失函数存在漏分割误分割等问题, 从表3中的评价指标上也能直观看出比本文方法低;

2) 组合损失函数能够克服单一损失函数的缺点, 能够利用不同损失函数的优点实现对烟尘图像进行更准确的分割. 从固定参数分别为($\beta = 0.6, \beta = 0.7, \beta = 0.8$)下的组合损失函数的实验结果可知, 组合损失函数能够利用CEL, SSIML和FL每个损失函数的优点, 能分割出相对更准确的目标烟尘区域, 但由于固定参数的组合损失不能很好的适应网络训练过程, 导致对于烟尘的边缘细节上的分割还不够准确, 如图5中(f)–(h)列和表3的评价指标中可以看出烟尘边缘分割不够精确;

3) 本文提出的动态加权组合损失函数能适应网络的训练过程, 在网络训练初期由CEL对烟尘目标整体区域进行分割, 提供先验指导信息, 随着网络训练的进程, 在分割出烟尘整体区域的基础上, 由SSIML和FL损失来关注烟尘边缘结构细节信息的提取, 使网络对于烟尘整体的分割完整度和边缘分割的准确性得到有效提升. 如图5中(i)–(k)列, 以及表3的评价指

标中可以看出烟尘分割结果更为准确. 从表3中的评价指标的对比得出, 在动态权重为(0.4, 0.8]范围内的IOU和DSC指标上最好, 在(0.3, 0.7]范围内的查全率R和F1度量指标上最好. 在本实验中, 更关注被分割出来的烟尘正样本数量, 更关注R和F1指标, 后续对比实验在动态权重为(0.3, 0.7]范围内进行. 实验结果表明, 在动态权重为(0.3, 0.7]时, 本文方法能有效增强网络对烟尘边缘的提取能力, 在R, IOU, DSC和F1这4个评价指标上比U-Net分别提高了9.25%, 5.27%, 4.31%和6.91%.

4.6 与其他模型对比

为了验证本文提出方法的有效性和适应性, 将本文方法C-UNet和其他11种语义分割方法进行对比, 对比方法有FCN^[26], U-Net^[15], ResNet^[27], PSPNet^[28], Deeplabv3+^[29], DA-Net^[30], W-UNet^[31], SE-DUNet^[32], Kidney-SegNet^[33], I-BiSeNet^[34], W-Net^[13]. 11种对比网络的训练参数与本文方法保持一致, 在6个烟尘场景测试集上进行实验, 评价指标对比如表4所示, 分割结果对比如图6所示.

表4 各场景下不同网络模型评价指标对比

Table 4 Comparison of evaluation indicators of different network models indifferent scenarios

场景	网络模型	R	P	IOU	DSC	F1
昏暗烟尘场景	FCN ^[26]	0.8263	0.8660	0.7407	0.8250	0.8340
	U-Net ^[15]	0.8635	0.8715	0.7749	0.8524	0.8651
	ResNet ^[27]	0.8242	0.8486	0.7273	0.8200	0.8290
	PSPNet ^[28]	0.8525	0.8613	0.7552	0.8387	0.8543
	Deeplabv3+ ^[29]	0.8719	0.8558	0.7610	0.8466	0.8686
	DA-Net ^[30]	0.8788	0.8479	0.7619	0.8450	0.8724
	W-UNet ^[31]	0.8660	0.8605	0.7608	0.8426	0.8649
	SE-DUNet ^[32]	0.9073	0.8543	0.7882	0.8673	0.8962
	Kidney-SegNet ^[33]	0.7760	0.7561	0.6172	0.7351	0.7719
	I-BiSeNet ^[34]	0.8746	0.8539	0.7684	0.8522	0.8704
	W-Net ^[13]	0.9098	0.8619	0.7915	0.8712	0.8998
	本文方法	0.9345	0.8609	0.8115	0.8844	0.9188
易辨烟尘场景	FCN ^[26]	0.8636	0.8360	0.7451	0.8356	0.8579
	U-Net ^[15]	0.8783	0.8301	0.7492	0.8428	0.8682
	ResNet ^[27]	0.8744	0.8136	0.7301	0.8353	0.8615
	PSPNet ^[28]	0.9008	0.8303	0.7543	0.8527	0.8858
	Deeplabv3+ ^[29]	0.9082	0.8197	0.7551	0.8534	0.8890
	DA-Net ^[30]	0.9088	0.8178	0.7548	0.8516	0.8890
	W-UNet ^[31]	0.9295	0.8224	0.7724	0.8645	0.9059
	SE-DUNet ^[32]	0.9414	0.8193	0.7761	0.8696	0.9142
	Kidney-SegNet ^[33]	0.8216	0.6857	0.5783	0.7180	0.7903
	I-BiSeNet ^[34]	0.9327	0.8240	0.7744	0.8675	0.9087
	W-Net ^[13]	0.9405	0.8204	0.7762	0.8690	0.9138
	本文方法	0.9642	0.8057	0.7796	0.8719	0.9277

(转下页)

(接上页)

干扰烟尘场景	FCN ^[26]	0.8097	0.8146	0.6878	0.7871	0.8107
	U-Net ^[15]	0.8880	0.7747	0.7009	0.8061	0.8628
	ResNet ^[27]	0.8196	0.7724	0.6726	0.7809	0.8097
	PSPNet ^[28]	0.8236	0.7876	0.6887	0.7916	0.8161
	Deeplabv3+ ^[29]	0.8539	0.7567	0.6648	0.7687	0.8325
	DA-Net ^[30]	0.8501	0.7606	0.6724	0.7748	0.8305
	W-UNet ^[31]	0.9042	0.7786	0.7140	0.8190	0.8759
	SE-DUNet ^[32]	0.9190	0.7513	0.7011	0.7970	0.8798
	Kidney-SegNet ^[33]	0.7507	0.6614	0.5224	0.6635	0.7309
	I-BiSeNet ^[34]	0.8624	0.7803	0.7075	0.8104	0.8446
	W-Net ^[13]	0.9299	0.7484	0.6975	0.7961	0.8869
	本文方法	0.9498	0.7525	0.7170	0.8120	0.9025
	小目标烟尘场景	FCN ^[26]	0.8033	0.7708	0.6590	0.7659
U-Net ^[15]		0.8430	0.7752	0.6797	0.7949	0.8286
ResNet ^[27]		0.8171	0.7596	0.6455	0.7764	0.8049
PSPNet ^[28]		0.8413	0.7886	0.6755	0.7960	0.8302
Deeplabv3+ ^[29]		0.8657	0.7780	0.6887	0.8074	0.8466
DA-Net ^[30]		0.8712	0.7464	0.6739	0.7914	0.8430
W-UNet ^[31]		0.8968	0.7647	0.7023	0.8147	0.8669
SE-DUNet ^[32]		0.9113	0.7581	0.7073	0.8202	0.8759
Kidney-SegNet ^[33]		0.8018	0.5920	0.5034	0.6473	0.7487
I-BiSeNet ^[34]		0.8889	0.7642	0.6927	0.8107	0.8608
W-Net ^[13]		0.9373	0.7533	0.7142	0.8277	0.8937
本文方法		0.9515	0.7514	0.7204	0.8341	0.9034
稀薄烟尘场景		FCN ^[26]	0.7977	0.8240	0.6633	0.7809
	U-Net ^[15]	0.8723	0.8072	0.7125	0.8204	0.8584
	ResNet ^[27]	0.8137	0.7651	0.6528	0.7706	0.8035
	PSPNet ^[28]	0.8403	0.7711	0.6828	0.7899	0.8256
	Deeplabv3+ ^[29]	0.8223	0.7999	0.6653	0.7835	0.8177
	DA-Net ^[30]	0.8868	0.7766	0.6976	0.8105	0.8623
	W-UNet ^[31]	0.8726	0.7842	0.6957	0.8104	0.8533
	SE-DUNet ^[32]	0.9239	0.7672	0.7152	0.8255	0.8877
	Kidney-SegNet ^[33]	0.7091	0.6751	0.4897	0.6411	0.7021
	I-BiSeNet ^[34]	0.8933	0.7809	0.7082	0.8177	0.8683
	W-Net ^[13]	0.9227	0.7745	0.7198	0.8265	0.8887
	本文方法	0.9445	0.7630	0.7266	0.8320	0.9016
	有颜色烟尘场景	FCN ^[26]	0.8984	0.8227	0.7467	0.8466
U-Net ^[15]		0.9057	0.7977	0.7293	0.8327	0.8818
ResNet ^[27]		0.8683	0.7875	0.7014	0.8143	0.8509
PSPNet ^[28]		0.9004	0.8058	0.7351	0.8383	0.8798
Deeplabv3+ ^[29]		0.9063	0.8014	0.7361	0.8361	0.8832
DA-Net ^[30]		0.9359	0.7860	0.7406	0.8434	0.9015
W-UNet ^[31]		0.9171	0.8072	0.7437	0.8459	0.8928
SE-DUNet ^[32]		0.9443	0.7925	0.7516	0.8501	0.9094
Kidney-SegNet ^[33]		0.8392	0.7300	0.6256	0.7587	0.8148
I-BiSeNet ^[34]		0.9215	0.7947	0.7402	0.8433	0.8930
W-Net ^[13]		0.9435	0.7918	0.7507	0.8488	0.9087
本文方法		0.9570	0.7863	0.7555	0.8518	0.9172

1) 在昏暗、易变、干扰、小目标、稀薄和有颜色烟尘场景下, FCN, U-Net, ResNet和PSPNet网络由于网

络结构设计相对简单, 对于非刚性多方向变化烟尘的整体和边缘细节信息提取能力欠佳, 导致分割的烟尘

存在漏分割和误分割的问题,如图6中的(c)–(f)列所示;

2) Deeplabv3+, DA-Net, W-UNet, SE-DUNet, Kidney-SegNet, I-BiSeNet和W-Net网络由于结合了更多的改进思路,分割效果有所提高,但是网络缺少针对性,网络的适应能力不足,仍然存在漏分割和误分割的情况,如图6中的(g)–(m)列所示,从表4中也能直观看出在不同烟尘场景下大部分的分割评价指标上都比本文方法低;

3) 本文方法能够利用CMDDD分解图像为包含烟尘主体信息的低频分量和包含烟尘边缘方向细节信息的8个高频分量,有效降低了烟尘细节信息的丢失,增强了网络对细节信息的提取能力,通过CDRU结合

跳跃连接的多方向细节信息,能有效提高网络对烟尘边缘方向细节信息的恢复能力,提高烟尘边缘细节部分的分割效果,并通过动态组合损失使网络更关注烟尘边缘信息的提取,本文方法的抗干扰能力得到了有效提高.在6个不同烟尘场景下不同方法的分割结果中,本文方法的分割结果优于FCN, U-Net, ResNet, PSPNet, Deeplabv3+, DA-Net, W-UNet, SE-DUNet, Kidney-SegNet, I-BiSeNet和W-Net共11种方法,分割烟尘整体及边缘更加准确.从表4的评价指标上也可清晰看出,在不同烟尘场景下,本文方法在R, IOU和F1这3个评价指标上都比其他分割模型表现好.通过实验证明,本文方法能够很好解决烟尘边缘信息丢失的问题,能适应复杂场景下目标烟尘的准确分割.

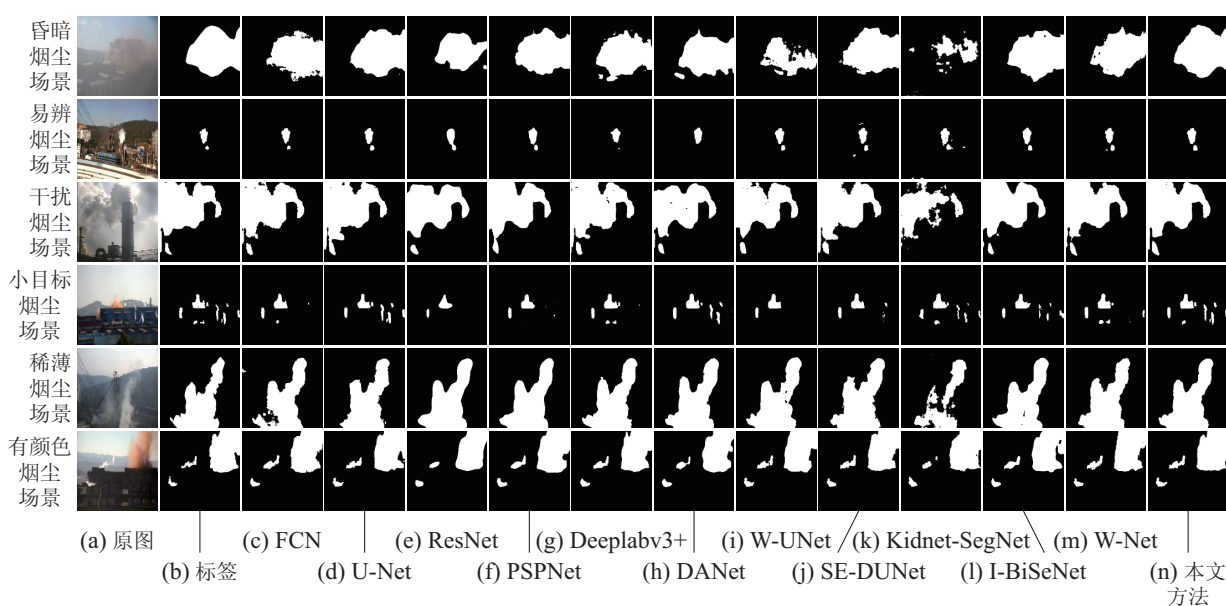


图6 不同烟尘场景实验结果对比

Fig. 6 Comparison of experimental results indifferent smoke scenes

5 结论

针对边缘模糊且方向多变烟尘分割不准确的问题,本文提出频域多方向C-UNet及动态损失的工业烟尘图像分割方法进行解决.首先利用CMDDD结构来提高网络对烟尘多方向细节特征的提取能力;然后通过高频多方向细节信息的跳跃连接,提升持续采样过程中细节信息的表达准确度;之后利用CDRU结构来增强烟尘边缘细节信息恢复能力;最后构建动态加权组合损失函数来优化训练网络,增强网络的自适应性和动态性,提高网络在训练过程中对烟尘边缘特征的提取能力.

实验结果表明,本文方法对边缘模糊方向多变烟尘有更好的分割效果,能针对性地改善烟尘边缘分割困难的问题.同时,为验证本文方法的适应性和有效性,分别在多个不同烟尘场景数据集下进行不同方法的对比实验,从实验结果中可知,本文方法的分割效

果都优于其他分割网络.

参考文献:

- [1] LIU Hui, WANG Yanan, CHEN Zongxin. Method to segment industrial emission smoke target based on adaptive variable step Gauss mixture model. *Control Theory & Applications*, 2020, 37(11): 2441 – 2450.
(刘辉, 王亚楠, 陈宗鑫. 自适应变步长高斯混合模型的工业烟尘目标分割方法. *控制理论与应用*, 2020, 37(11): 2441 – 2450.)
- [2] LIU Hui, ZHANG Junpeng, LI Qingrong. Industrial smoke target segmentation based on fully convolutional networks with multiscale convolution and dynamic weight loss function. *Journal of Computer-Aided Design & Computer Graphics*, 2020, 32(12): 1898 – 1909.
(刘辉, 张俊鹏, 李清荣. 多尺度卷积与动态权重代价函数的全卷积网络工业烟尘目标分割. *计算机辅助设计与图形学学报*, 2020, 32(12): 1898 – 1909.)
- [3] ZHAN Guangli, LIU Hui, YANG Lu. Industrial smoke image segmentation based on improved attention W-Net. *Computer Integrated Manufacturing System*, 2023, 29(2): 628 – 637.
(詹光莉, 刘辉, 杨路. 基于改进注意力W-Net的工业烟尘图像分割.

- 计算机集成制造系统, 2023, 29(2): 628 – 637.)
- [4] WANG X, JIANG A, WANG Y. A segmentation method of smoke in forest-fire image based on fbm and region growing. *The Fourth International Workshop on Chaos-Fractals Theories and Applications*. Hangzhou: IEEE, 2011: 390 – 393.
- [5] YUAN F. Video-based smoke detection with histogram sequence of LBP and LBPV pyramids. *Fire Safety Journal*, 2011, 46(3): 132 – 139.
- [6] HSU Y C, DILLE P, SARGENT R, et al. Industrial smoke detection and visualization. *ArXiv Preprint*, 2018, arXiv.1809.06263
- [7] CALDERARA S, PICCININI P, CUCCHIARA R. Smoke detection in video surveillance: A MoG model in the wavelet domain. *International Conference on Computer Vision Systems*. Santorini: Springer-Verlag, 2008: 373 – 379.
- [8] GENG J, LIU J, KONG X, et al. The fishmeal adulteration identification based on microscopic image and deep learning. *Computers and Electronics in Agriculture*, 2022, 198: 106974.
- [9] TAN L, LV X, LIAN X, et al. YOLOv4-Drone: UAV image target detection based on an improved YOLOv4 algorithm. *Computers & Electrical Engineering*, 2021, 93: 107261.
- [10] MO Y, WU Y, YANG X, et al. Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing*, 2022, 493: 626 – 646.
- [11] YUAN F, ZHANG L, XIA X, et al. Deep smoke segmentation. *Neurocomputing*, 2019, 357: 248 – 260.
- [12] DONG Zeshu, YUAN Feiniu, XIA Xue. Improved spatial and channel information based global smoke attention network. *Journal of Beijing University of Aeronautics and Astronautics*, 2022, 48(8): 1471 – 1479.
(董泽舒, 袁非牛, 夏雪. 基于改进空间通道信息的全局烟雾神经网络. 北京航空航天大学学报, 2022, 48(8): 1471 – 1479.)
- [13] ZHAN Guangli, LIU Hui, CHEN Fugang, et al. Connection attention mechanism and residual ASPP of W-Net industrial smoke image segmentation. *Control Theory & Applications*, 2023, 40(1): 160 – 171.
(詹光莉, 刘辉, 陈甫刚, 等. 衔接注意力机制与残差ASPP的W-Net工业烟尘图像分割. 控制理论与应用, 2023, 40(1): 160 – 171.)
- [14] 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.
(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.)
- [15] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer, 2015: 234 – 241.
- [16] YANG Y, FENG C, WANG R. Automatic segmentation model combining U-Net and level set method for medical images. *Expert Systems with Applications*, 2020, 153: 113419.
- [17] SAMBYAL N, SAINI P, SYAL R, et al. Modified U-Net architecture for semantic segmentation of diabetic retinopathy images. *Biocybernetics and Biomedical Engineering*, 2020, 40(3): 1094 – 109.
- [18] ZHANG S, ZHANG C. Modified U-Net for plant diseased leaf image segmentation. *Computers and Electronics in Agriculture*, 2023, 204: 107511.
- [19] DO M N, VETTERLI M. The contourlet transform: A efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 2005, 14(12): 2091 – 2106.
- [20] BURT P J, ADELSON E H. The Laplacian pyramid as a compact image code. *Readings in Computer Vision*. Morgan Kaufmann: Elsevier, 1987: 671 – 679.
- [21] BAMBERGER R H, SMITH M J T. A filter bank for the directional decomposition of images: Theory and design. *IEEE Transactions on Signal Processing*, 1992, 40(4): 882 – 893.
- [22] CHEN Daqian, ZHANG Fan, HAO Pengyi, et al. 2D medical image segmentation combining multi-Scale channel attention and boundary enhancement. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(11): 1742 – 1752.
(陈大千, 张凡, 郝鹏翼, 等. 结合多尺度通道注意力和边界增强的2D医学图像分割. 计算机辅助设计与图形学学报, 2022, 34(11): 1742 – 1752.)
- [23] QIN X, ZHANG Z, HUANG C, et al. BASNet: Boundary-aware salient object detection. *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, 2019: 7471 – 7481.
- [24] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318 – 27.
- [25] BOUGOURZI F, DORNAIKA F, TALEB-AHMED A. Deep learning based face beauty prediction via dynamic robust losses and ensemble regression. *Knowledge-Based Systems*, 2022, 242: 108246.
- [26] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE, 2015: 3431 – 3440.
- [27] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 770 – 778.
- [28] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 6230 – 6239.
- [29] 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: Springer, 2018: 801 – 818.
- [30] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE, 2019, 3141 – 3149.
- [31] LI Q, SHEN L. WaveSNet: Wavelet integrated deep networks for image segmentation. *Pattern Recognition and Computer Vision: The 5th Chinese Conference, PRCV 2022*. Shenzhen, China: Springer Nature Switzerland, 2022: 325 – 337.
- [32] CHENG J, TIAN S, YU L, et al. A deep learning algorithm using contrast-enhanced computed tomography (CT) images for segmentation and rapid automatic detection of aortic dissection. *Biomedical Signal Processing and Control*, 2020, 62: 102145.
- [33] AATRESH A A, YATGIRI R P, CHANCHAL A K, et al. Efficient deep learning architecture with dimensionwise pyramid pooling for nuclei segmentation of histopathology images. *Computerized Medical Imaging and Graphics*, 2021, 93: 101975.
- [34] LI Y, ZHANG W, LIU Y, et al. A lightweight network for real-time smoke semantic segmentation based on dual paths. *Neurocomputing*, 2022, 501: 258 – 69.

作者简介:

张大锦 硕士研究生, 目前研究方向为图像处理、机器学习、数据挖掘等, E-mail: 2328668736@qq.com;

刘辉 教授, 硕士生导师, 目前研究方向为机器学习、数据挖掘、图像处理等, E-mail: liuhui621@126.com;

陈甫刚 工程师, 目前研究方向为数据挖掘、图像处理、工业信息处理等, E-mail: 357319154@qq.com;

赵安 硕士研究生, 目前研究方向为机器学习、数据挖掘、图像处理等, E-mail: 1304134016@qq.com.