

基于 DeepLabv3+网络的电流互感器红外图像分割方法

袁刚¹, 许志浩¹, 康兵¹, 罗吕¹, 张文华¹, 赵天成²

(1. 南昌工程学院 电气工程学院, 江西 南昌 330099; 2. 国网吉林省电力有限公司电力科学研究院, 吉林 长春 130021)

摘要: 红外图像智能分析是变电设备故障诊断的一种有效方法, 目标设备分割是其关键技术。本文针对复杂背景下电流互感器整体分割难的问题, 采用基于 ResNet50 的 DeepLabv3+神经网络, 用电流互感器的红外图像训练语义分割模型的方法, 对收集到的样本采用限制对比度自适应直方图均衡化方法实现图像轮廓增强, 构建样本数据集, 并运用图像变换扩充样本数据集, 搭建语义分割网络训练语义分割模型, 实现电流互感器像素与背景像素的二分类。通过文中方法对 420 张电流互感器红外图像测试, 结果表明, 该方法的平均交并比 (Mean Intersection over Union, MIoU) 为 87.5%, 能够从测试图像中精确分割出电流互感器设备, 为后续电流互感器的故障智能诊断做铺垫。

关键词: 红外图像; 电流互感器; ResNet50; DeepLabv3+; 语义分割

中图分类号: TN219; TM452 文献标志码: A 文章编号: 1001-8891(2021)11-1127-08

DeepLabv3+ Network-based Infrared Image Segmentation Method for Current Transformer

YUAN Gang¹, XU Zhihao¹, KANG Bing¹, LUO Lyu¹, ZHANG Wenhua¹, ZHAO Tiancheng²

(1. School of Electrical Engineering Nanchang Institute of Technology, Nanchang 330099, China;

2. Electric Power Science Research Institute of State Grid Jilin Electric Power Co., LTD., Changchun 130021, China)

Abstract: Infrared image intelligent analysis is an effective method for the fault diagnosis of transformer equipment, and its key technology is target device segmentation. In this study, aiming to address the difficulty in overall segmentation of current transformers with complex backgrounds, the DeepLabv3+ neural network based on ResNet50 was applied to train the semantic segmentation model with infrared image of CT. The collected samples were enhanced by the limited contrast adaptive histogram equalization method, and a sample dataset was constructed. The sample dataset was expanded by image distortion, and a semantic segmentation network was built to train the semantic segmentation model to realize the binary classification of current transformer pixels and background pixels. The test results of 420 current transformer infrared images showed that the MIOU of this method is 87.5%, which can accurately divide the current transformer equipment from the test images and lay a foundation for the subsequent intelligent fault diagnosis of current transformers.

Key words: infrared image, current transformer, ResNet50, DeepLabv3+, semantic segmentation

0 引言

在变电设备红外图像分析的过程中, 高质量的图像能简化故障诊断算法流程。然而, 真实的变电站场景复杂, 背景干扰大, 拍摄的红外图像成像质量低; 此外, 某些红外热像仪还自带标记, 红外图像中的比色条、热像仪 logo、温度标签等覆盖了图像中部分温度信息, 造成数据污染, 增加了变电设备故障诊断难

度^[1]。当前基于红外图像变电设备故障自动诊断主要分为 3 步: 查找感兴趣区域 (Region of Interest, ROI)、特征提取和状态分类^[2]。其中查找 ROI 的目的主要是缩小目标对象分析范围, 主流方法包括目标分割与目标检测两种手段。该类方法在实施过程中可分为两种思路, 即一步定位热点区域法和多步定位热点区域法。

一步定位热点区域是用一些分割算法或目标检

收稿日期: 2021-08-02; 修订日期: 2021-10-16.

作者简介: 袁刚 (1997-), 男, 贵州盘州人, 硕士研究生, 研究方向为电力设备故障检测与诊断。E-mail: 862635457@qq.com.

通信作者: 许志浩 (1988-), 男, 湖北武汉人, 讲师, 博士, 硕导, 研究方向为电力设备智能检测与人工智能应用。E-mail: zhxuhi@whu.edu.cn.

基金项目: 吉林省电力科学研究院有限公司科技项目 (KY-GS-20-01-07)。

测算法直接从变电设备的红外图像中分割或界定过热区域。康龙等^[3]利用红外图像灰度直方图确定聚类中心和聚类个数，用遗传算法来确定最优聚类中心，最后用模糊 C 均值 (Fuzzy C-means) 来分割过热区域；曾亮等^[4]用大津 (OTSU) 算法和区域生长法分割过热区域；Hui Zou 等^[5]利用 k 均值聚类算法 (k-means clustering algorithm) 将灰度图像分割成 k 个区域，计算 k 个区域中的平均值，将其作为聚类中心不断迭代来分割最终的过热区域。随着深度卷积网络在图像处理领域取得突破性进展，目标检测算法定位故障区域的方式也取得了较好的成绩。王旭红等^[6]利用 Single Shot MultiBox Detector (SSD)，林颖等^[7]利用 You Only Look Once (YOLO)，刘云鹏等^[8]利用 Faster Region-based CNN 等目标检测网络实现红外输变电设备异常发热区域界定。

然而，上述方法难以自动规避数据污染和背景热源干扰，处理结果都缺少对象语义表达，且主要适用于发热较大的电流致热型故障；而对于发热较小的电压致热产生的热点区域却难以做到精确定位，甚至失败。此外，用神经网络直接实现变电设备故障区域界定的算法需要大量的训练数据集，红外故障数据集样本量不足是限制该方法的主要问题。

综上所述，用一步定位热点区域方法实现局部过热区域定位难以满足变电设备故障检测需求，利用多步定位热点区域可以提高故障分析准确率。这类方法的主要思想是：先将目标设备整体从复杂背景中进行分割，减少数据污染和背景干扰，再进一步分割可疑热区域，保证后续提取的热特征量的空间位置信息是来自被分析的变电设备本体^[1]。

图像语义分割是一种能让计算机理解图像的技术，在一个网络中同时实现图像中的物体类别识别和高精度的图像分割。近年来，该技术在可见光的图像语义分割已经取得不俗的效果^[9]，而对变电设备红外图像的场理解的研究还处于初级阶段。本文用电流互感器作为训练样本，DeepLabv3+网络训练模型^[10]，从复杂背景中分割设备整体，结合图像形态学方法对语义分割结果进行后处理，以提高最终分割的精度，为后续变电设备不同致热因素导致的故障预判做铺垫。

1 算法理论

语义分割技术是利用计算机将图像中的像素按照图像表达的语义信息进行分类^[10]。在以往的分类网络中，会将图片进行降维处理，丢弃图片原有的空间信息。而图像的语义分割算法需要分类每个像素并将分类结果还原成带有语义信息的与原图同等大小的

图片。因此，保留像素的空间信息对于图像语义分割而言尤为重要。随着全卷积神经网络 (Fully Convolutional Networks, FCN) 的提出，Long^[11]等人将 AlexNet, VGGNet 和 GoogLeNet 等主流分类网络中的全连接层替换成卷积层，并在最后添加转置卷积将特征图恢复到原来的尺寸^[12]，图像像素的空间位置特征得以保留，语义分割才有了突破性发展。本文利用近年来在公共数据集中测试 MIoU 得分最高的 DeepLabv3+ 作为训练变电设备语义分割模型的网络。

1.1 DeepLabv3+网络结构

编码解码 (Encode-Decode) 结构^[13]是语义分割网络中的主流结构，所谓的编码过程是通过特征提取网络提取变电设备的特征，再经过解码实现特征信息重组，在这个过程中，网络根据图像的标签信息不断修正参数，最终实现监督式学习的对像语义分割。

DeepLabv3+的特征提取主干网络为残差神经网络 ResNet^[14]。深度残差网络的设计是为了克服由于网络深度加深而产生的学习效率变低与准确率无法有效提升的问题。该网络允许一部分输入不经过卷积网络传到输出，保留了部分浅层信息，避免了因特征提取网络的加深而导致特征细节的丢失^[15]，残差模块的引入可以使网络在加深的同时保证较高的准确率，且更易于网络的优化，提高网络收敛速度。

ResNet 残差网络原型如图 1(a)所示，网络结构主要由卷积残差块 (Conv-block) 和恒等残差块 (Identity-block) 两部分基本块组成，如图 1(b)(c)所示。以 ResNet50 为例，其中 Conv-block 共 4 个，是用来调整输入的长、宽及通道数；Identity-block 共 12 个，是用来串联网络，增加网络层数。残差网络的计算公式如下：

$$y_i = H(x_i) + F(x_i, W_i) \quad (1)$$

$$x_{i+1} = f(y_i) \quad (2)$$

式(1)、(2)中： x_i , y_i 分别表示第 i 个残差网络的输入和输出； x_{i+1} 表示第 $i+1$ 层残差网络的输入； W_i 表示卷积操作； $F(x_i, W_i)$ 表示残差函数； f 表示 Relu 激活函数。其中 $H(x_i)$ 表示残差边的输出，在 Conv-block 网络中残差边的输出需要调整输入图像的大小并进行归一化，而在 Identity-block 结构中残差边的输入和输出恒等，这两种网络结构输出计算公式为：

$$H(x_i) = \begin{cases} \frac{\sum_{m=-\infty}^{\infty} x_i g(m, n)}{255} \\ x_i \end{cases} \quad (3)$$

式中： $g(m, n)$ 为卷积核。

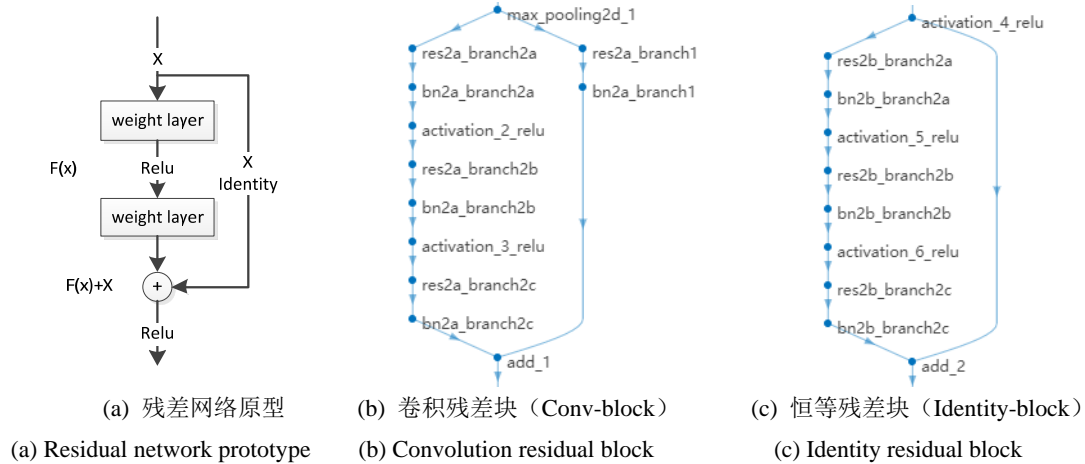


图1 Conv Block 与 Identity Block 的结构

Fig.1 Structure of Conv Block and Identity Block

在整个编码网络结构中为了获得更加高级的图像特征，在 ResNet50 网络末端添加空洞卷积金字塔模块 (Atrous Spatial Pyramid Pooling, ASPP)^[16]，它接受基网络提取的高级特征图作为输入，经过具有不同空洞速率的空洞卷积，最后将结果输出进行堆叠，达到覆盖多尺度感受野的目的。

上述两部分构成了 DeepLabv3+ 网络的编码模块。将编码模块输出的特征图与编码模块中的浅层特征相结合，经过上采样步骤构成解码模块，该模块的功能是利用转置卷积将特征图像进行逐层放大最终

还原成与输入图像尺寸同等大小的图像，最后经过 Softmax Layer 求出输出图像的像素所在类别的概率，Pixel Classification Layer 对像素进行分类^[17]，输出网络预测的 label。网络结构如图 2 所示。其中 softmax 函数定义为：

$$p_k(z_i) = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}} \quad (4)$$

式中： z_i 表示第 i 个节点输出的像素值； K 表示网络分类的分类个数； $p_k(z_i)$ 表示 z_i 属于第 k 个类别的概率。

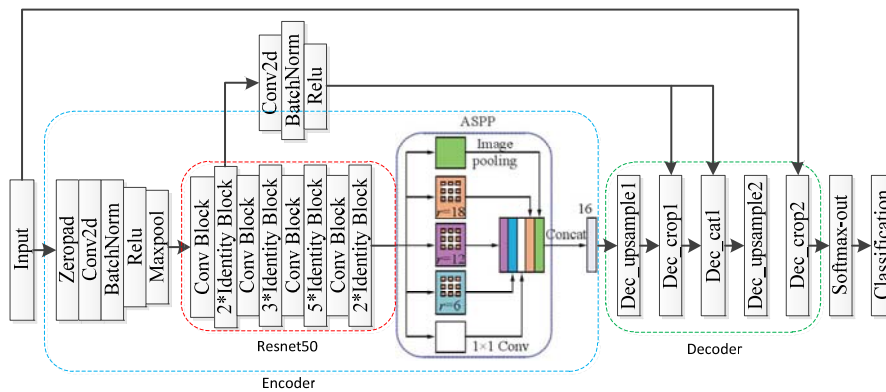


图2 DeepLabv3+结构图

Fig.2 DeepLabv3+ structure diagram

1.2 数学形态学运算

考虑到训练的模型难以准确无误地完成对输入图像像素分类，或多或少会有一些与设备本体颜色相近的像素点被分割出来。因此本文采用数学形态学中的开闭运算对分割后的图片进行处理。利用开运算对分割后的图像进行腐蚀，之后做膨胀运算，平滑分割后图像的轮廓，同时删除一些误分割的像素块，该运算的数学表达式为：

$$A \circ B = (A \ominus B) \oplus B \quad (5)$$

利用闭运算对图像做膨胀之后再行腐蚀处理，填充一些图像中出现的孔洞或裂口，其表达式为：

$$A \bullet B = (A \oplus B) \ominus B \quad (6)$$

以上两式中 A 为待处理图像； B 是单个结构化元素对象。

2 电流互感器红外图像数据集预处理

由于拍摄设备红外图像的热像仪型号各异，拍摄环境复杂，部分热摄像头拍摄的图片噪点多，此外，

当有设备表面温度与环境温度相差较小时，设备特征被淹没在环境中，丢失了大量细节特征，不利于图像标注和网络提取设备的轮廓特征。因此需要对收集到的图像做增强处理。

2.1 基于对比度限制自适应直方图均衡化的图像增强算法

由于红外图像反应的是设备温度和环境温度的高低分布，当环境温度与设备本体温度相近时或在拍摄时设置较大的温度区间，将使设备轮廓与背景环境混为一体，边界模糊。如采用直方图均化（Histogram Equalization, HE）方法对全局图像进行增强，直方图大峰值可能落在背景噪声或非感兴趣区域；在这种情况下，直方图均衡化会导致图像背景和设备灰度值提高，对比度降低。处理效果图和直方图如图 3(b), (e) 所示。

与上述方法相比，使用自适应直方图均衡化（Adaptive histogram equalization, AHE）^[18]改进的对比度限制自适应直方图均衡化（Contrast limit adaptive histogram equalization, CLAHE）^[19]算法能够在一定程度上抑制噪声的放大，这主要是通过限制 AHE 算法的对比提高程度来达到的。CLAHE 通过在计算累积直方图函数前用预先定义的阈值来裁剪直方图以达到限制放大幅度的目的。在处理过程中，对于给定的输入图像，CLAHE 算法将图像分割为互不重叠的图像块，对划分后的每一个子块计算其对应的直方图，使用预先设定的阈值 L 对每个子块直方图进行裁剪，同时统计整个直方图中超过上限阈值 L 的像素数，并将这些像素数重新分布到对应子块的直方图中。最后，通过使用双线性插值来消除边界伪影，在子矩阵上下文区域内实现像素的新灰度级分配的计算。处理效果图和直方图如图 3(c), (f) 所示。

2.2 数据集制作与划分

在制作训练数据集时，使用图像标注软件将上一步预处理好的图像中占据主体像素的设备进行精细标注，其余区域都视作背景。此外电流互感器的常见故障为电流致热型故障，常发生于导线与设备的连接处，在标注时应把导线与设备连接处也视为设备的一个特征包含到设备本体当中，如图 4 所示。



(a) 原图 (b) HE 算法处理图像 (c) CLAHE 算法处理图像
(a) Original image (b) HE algorithm processes image (c) Image processing by CLAHE algorithm

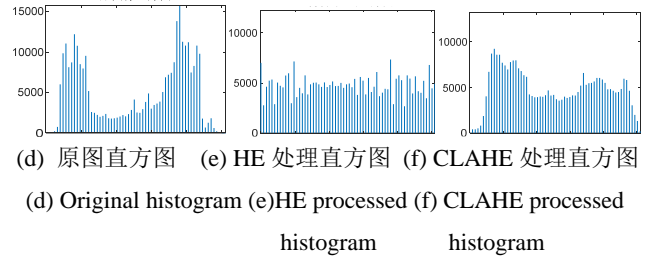


图 3 图像增强及对应直方图

Fig.3 Image enhancement and corresponding histogram



(a) 原图
(a) Original image



(b) 标签图
(b) Label image

图 4 数据集中原图与标签图

Fig.4 Original image and label image in dataset

数据扩充是一种常见的技术，已被证明有利于机器学习模型的一般培训，特别是深层架构，要么加速收敛，要么充当正则化器，从而避免过拟合，提高泛化能力^[20]。数据量的不足是当前网络模型泛化能力弱的关键因素。本文收集到的电流互感器红外图像数量有限，共 700 张，在训练时还要从中划分一部分作为验证集和测试集，以这些数据来训练语义分割模型难以达到一个好的分割精度。因此，采取图像扭曲变换来创建新样本扩充数据集，以避免训练过程中出现过拟合。针对电流互感器的红外图像数据集采用平移、旋转、翻转等图像变换方式来扩充，如图 5 所示，每一张图片和对应的标签经过 3 次变换，最终样本总数为 2100 张。将上述处理好的数据集划分 60% 作为训练样本集，用来建立像素分类器模型；划分 20% 作为验证集，用来验证训练效果并且做超参数调整；划分 20% 作为测试集，用来测试模型的泛化能力。

3 实验过程与分析

3.1 实验流程

本文中通过构建电流互感器红外图像作为训练语义分割模型的数据集，该数据集中包含了多种不同电压等级的电流互感器，同时也包含了不同拍照角

度,不同背景和不同气候条件下的图像,这些图像能够真实反映变电站内的复杂环境。训练模型的网络是以 ResNet50 为基网络的 DeepLabv3+。此外,本文还设计了多组对比实验,第一组是基于 ResNet50 的 DeepLabv3+模型和基于 ResNet18 的 DeepLabv3+模型对比;第二组是用上述两个模型和 FCN-8s、SegNet 模型对比;第三组是在 DeepLabv3+ (基网络为 ResNet50) 网络之后加入形态学开闭运算对比。

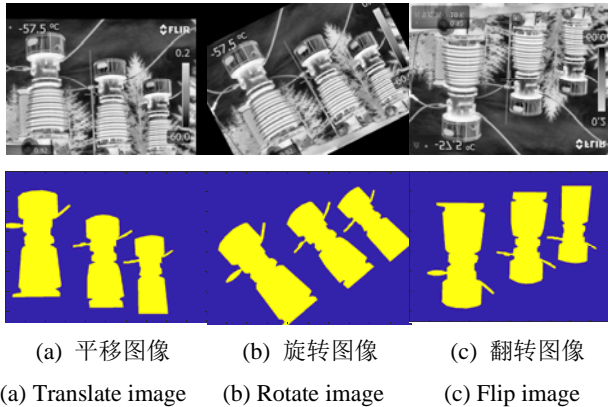


图5 原图像数据集与扩充数据

Fig.5 Original image dataset and extended data

由于实验过程需要进行大量的图像计算,对电脑的内存和 GPU 性能要求都比较高。因此本文在工作站平台上以 Matlab 的深度学习工具箱作为实验软件平台,搭建了 FCN-8s, SegNet, DeepLabv3+ (基网

络为 ResNet50)和 DeepLabv3+ (基网络为 ResNet18) 网络进行分别训练。电脑的硬件参数为: CPU 为 intel(R)Xeon(R) Gold5120T, 128 G 内存, 配备 QuadorP2000 显卡。

在训练过程中,首先对数据集进行扩充,然后将数据集分批送入网络训练。设置训练最小批次为 10, 迭代次数为 9800 次,学习率采用分段调整,设置初始学习率为 0.01,每迭代 10 轮低 0.1,这允许网络以更高的初始学习率快速学习,而一旦学习率下降,能够求得接近局部最优的解。通过设置 'Validation Data' 参数,在每轮都对验证数据对网络进行测试。'Validation Patience' 设置为 4,在验证准确度收敛时提前停止训练,这可以防止网络对训练数据集进行过拟合。采用带动量的随机梯度下降 (Stochastic Gradient Descent with Momentum, SGDM) 算法,动量参数为 0.9。利用交叉熵计算预测误差,进行反向传递,更新网络参数。损失函数为:

$$E = \sum_k^K w_k \log(p_k(z_i)) \quad (7)$$

式中: w_k 为类别 k 的损失权重; $p_k(z_i)$ 为像素 z_i 属于真实类别 k 的概率。训练过程如图 6 所示,随着学习率的下降,最终训练 6850 次,验证准确率为 95.49%,验证损失为 0.1189。

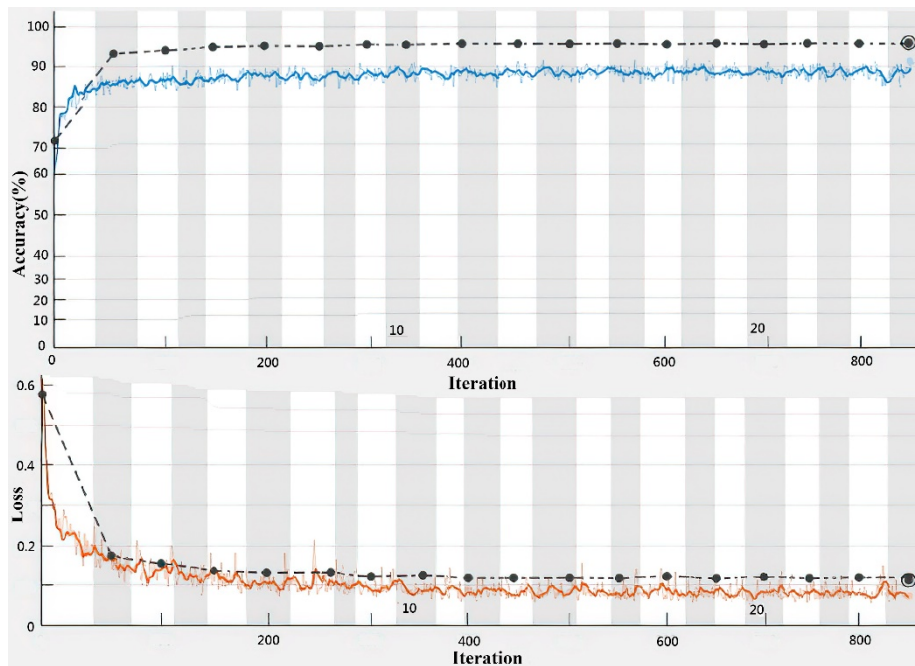


图6 基于 Resnet50 的 DeepLabv3+网络模型训练过程

Fig.6 Training process of DeepLabv3+ network model based on ResNet50

3.2 评估指标

在实验中用 $MIoU^{[21]}$ 作为分割结果评价指标,它计算两个集合的交集和并集之间的比值的平均值,可

以反应分割结果和真实标签的重合程度。假设 k 为像素分类类别总数, p_{ij} 和 p_{ji} 分别表示预测结果为 i 实际结果为 j 的像素总数和预测结果为 j , 实际结果为 i 的

像素总数，而 p_{ii} 表示预测结果为 i ，真实结果也为 i 的像素总数。计算公式如(8)所示。

$$MIoU = \frac{1}{k+1} \frac{\sum_{i=0}^k p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (8)$$

3.3 测试数据集分析

表 1 是基于 ResNet50 的 DeepLabv3+，基于 ResNet18 的 DeepLabv3+，SegNet 和 FCN-8s 等模型在 420 张测试数据集上测试的结果。

表 1 多种模型测试数据表

Table 1 Test data table of various models

Model	Categories	Accuracy	IoU	MIoU
DeepLabv3+ (ResNet50)	CT	0.86	0.77	0.855
	Background	0.95	0.94	
DeepLabv3+ (ResNet18)	CT	0.81	0.72	0.81
	Background	0.92	0.90	
SegNet	CT	0.67	0.44	0.615
	Background	0.86	0.79	
FCN-8s	CT	0.75	0.63	0.74
	Background	0.89	0.85	

在测试过程中数据集中依然使用多种不同电压等级的电流互感器进行测试，测试结果如图 7 所示，其中(a)表示测试图，(b)表示测试图的标签图，(c)表示基于 ResNet50 的 DeepLabv3+模型的预测图，(d)表示基于 ResNet18 的 DeepLabv3+模型的预测图，(e)表示 SegNet 模型的预测图，(f)表示 FCN-8s 模型的预测图。从实际的测试结果可以看出，4 种模型都能够从复杂背景中预测设备空间位置和部分轮廓，但是依然难以避免地会出现部分不属于电流互感器的像素被分类到该类别中，属于电流互感器类别的像素却被预测成背景，其中 SegNet 模型误分割和漏分割程度更高，FCN-8s 模型次之，与 DeepLabv3+的两个模型相比，SegNet 和 FCN-8s 模型分割结果较为粗糙；而基于 ResNet18 的 DeepLabv3+模型的分割结果能较大程度上接近原标签图像，但与基于 ResNet50 的 DeepLabv3+模型相比而言，基于 ResNet50 的 DeepLabv3+模型对图像分割的细节控制更好。

对比实验结果可以看出使用残差网络和 ASPP 模块的 DeepLabv3+相比于以 VGG-16 为特征提取网络的 SegNet 和 FCN-8s 网络结构能够提取更加高级的特征，同时 DeepLabv3+在进行上采样时能够融合大量的浅层信息特征，对于后期的像素分类和分割结果能够保留更多细节。由于本文用于训练的图像数量小，因此，选择合适的网络结构和网络可训练参数量是训练模型的关键。经本文实验得出，基于 ResNet50 的 DeepLabv3+网络能够满足在小样本条件下训练语义分割模型从复杂背景环境下分割电流互感器设备的

要求。



(a)测试图

(a) Test image



(b) 原图标签图

(b) Original label image



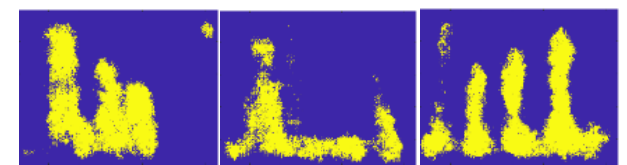
(c) DeepLabv3+ (ResNet50) 模型分割图

(c) DeepLabv3+ (ResNet50) model segmentation image



(d) DeepLabv3+ (ResNet18) 模型分割图

(d) DeepLabv3+ (ResNet18) model segmentation image



(e) SegNet 模型分割图

(e) SegNet model segmentation image



(f) FCN-8s 模型分割图

(f) FCN-8s model segmentation image

图 7 分割图像

Fig.7 Model segmentation and post-processing image

对于基于 ResNet50 的 DeepLabv3+模型误分割的情况，采用形态学中的开运算去除误分割产生的小区域像素，同时处理分割边缘，使电流互感器边缘变得平滑。根据统计，单个电流互感器在图像中的面积均大于 1200，而误分割区域通常较小，因此可以设置一个面积阈值为 1200，当面积小于 1200 时删除该区

域,其余像素保留;开运算操作后进行闭运算,使图像中的孔洞封闭。加入数学形态学运算处理后在测试数据集中的表现如表2所示,处理图像如图8所示,其中(a)表示基于 ResNet50 的 DeepLabv3+模型的预测图,(b)表示预测图经过后处理的图像。

表2 基于 ResNet50 的 DeepLabv3+模型加入后处理前后测试对比

Table 2 Comparison of tests before and after the addition of the DeepLabv3+ model based on ResNet50

Model	Categories	Accuracy	IouU	MIoU
DeepLabv3+ (ResNet50)	CT	0.86	0.77	0.855
	Background	0.95	0.94	
Our algorithm	CT	0.87	0.79	0.875
	Background	0.97	0.96	



(a) DeepLabv3+ (ResNet50) 模型分割图

(a) DeepLabv3+ (ResNet50) model segmentation image



(b) 后处理图像

(b) Post-processing image

图8 语义分割后处理图像

Fig.8 Semantic segmentation after image processing

4 总结

本文以电流互感器为研究对象,针对复杂背景和图片中数据污染的电流互感器设备分割问题,采用了基于 ResNet50 的 DeepLabv3+神经网络,利用电流互感器红外图像训练语义分割模型。通过实验表明:基于 ResNet50 的 DeepLabv3+神经网络训练的语义分割模型能够较为精细地从红外图像中分割出电流互感器,并且结合图像的开闭运算能够处理误分割的像素点,最终在测试集上的 MIoU 为 0.875,有效地提高了分割精度,最终实现电流互感器设备像素与背景像素的分类。

限于文章篇幅有限,本文仅研究了深度学习方法用于电流互感器红外图像场景理解的问题,仅是分割出目标设备。在后续的工作当中,逐步实现可疑故障区域的分割以及故障类型分类,最终实现电流互感器

设备故障自动诊断。

参考文献:

- [1] 王小芳,毛华敏.一种复杂背景下的电力设备红外图像分割方法[J].红外技术,2019,41(12):1111-1116.
WANG Xiaofang, MAO Huamin. Infrared Image Segmentation Method for Power Equipment in Complex Background[J]. *Infrared Technology*, 2019, 41(12): 1111-1116.
- [2] GONG X, YAO Q, WANG M, et al. A deep learning approach for oriented electrical equipment detection in thermal images[J]. *IEEE Access*, 2018: 1-1. Doi: 10.1109/ACCESS.2018.2859048.
- [3] 康龙.基于红外图像处理的变电站设备故障诊断[D].北京:华北电力大学,2016.
KANG Long. Substation equipment fault diagnosis based on infrared image processing[D]. Beijing: North China Electric Power University, 2016.
- [4] 曾亮.基于红外图像的变电站设备故障精准定位方法的研究[D].重庆:重庆理工大学,2019.
ZENG Liang. Research on precise fault location method of substation equipment based on infrared image[D]. Chongqing: Chongqing University of Technology, 2019.
- [5] ZOU H, HUANG F. A novel intelligent fault diagnosis method for electrical equipment using infrared thermography[J]. *Infrared Physics & Technology*, 2015, 73: 29-35.
- [6] 王旭红,李浩,樊绍胜,等.基于改进 SSD 的电力设备红外图像异常自动检测方法[J].电工技术学报,2020,35(S1):302-310.
WANG Xuhong, LI Hao, FAN Shaosheng, et al. Infrared image anomaly automatic detection method for power equipment based on improved single shot multi box detection[J]. *Transactions of China Electrotechnical Society*, 2020, 35(S1): 302-310.
- [7] 林颖,郭志红,陈玉峰.基于卷积递归网络的电流互感器红外故障图像诊断[J].电力系统保护与控制,2015,43(16):87-94.
LIN Ying, GUO Zhihong, CHEN Yufeng. Convolutional-recursive network based current transformer infrared fault image diagnosis[J]. *Power System Protection and Control*, 2015, 43(16): 87-94.
- [8] 刘云鹏,裴少通,武建华,等.基于深度学习的输变电设备异常发热红外图片目标检测方法[J].南方电网技术,2019,13(2):27-33.
LIU Yunpeng, PEI Shaotong, WU Jianhua, et al. Deep learning based target detection method for abnormal hot spots infrared images of transmission and transformation equipment[J]. *Southern Power System Technology*, 2019, 13(2): 27-33.
- [9] 王晨.基于深度学习的红外图像语义分割技术研究[D].上海:中国科学院大学(中国科学院上海技术物理研究所),2017.
WANG Chen. Research on infrared image semantic segmentation technology based on deep learning[D]. Shanghai: University of Chinese

- Academy of Sciences (Shanghai Institute of Technical Physics, Chinese Academy of Sciences), 2017.
- [10] 邝辉宇, 吴俊君. 基于深度学习的图像语义分割技术研究综述[J]. 计算机工程与应用, 2019, 55(19): 12-21, 42.
- KUANG Huiyu, WU Junjun. Survey of image semantic segmentation based on deep learning[J]. *Computer Engineering and Applications*, 2019, 55(19): 12-21, 42.
- [11] LONG J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431-3440.
- [12] 袁铭阳, 黄宏博, 周长胜. 全监督学习的图像语义分割方法研究进展[J]. 计算机工程与应用, 2021, 57(4): 43-54.
- YUAN Mingyang, HUANG Hongbo, ZHOU Changsheng. Research progress of image semantic segmentation based on fully supervised learning[J]. *Computer Engineering and Applications*, 2021, 57(4): 43-54.
- [13] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[C]//*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495, DOI: 10.1109/TPAMI.2016.2644615.
- [14] Garcia-Garcia A, Orts-Escobedo S, Oprea S, et al. A review on deep learning techniques applied to semantic segmentation[J/OL]. *Computer Vision and Pattern Recognition*, 2017. <https://arxiv.org/abs/1704.06857>.
- [15] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[J/OL]. *Computer Vision and Pattern Recognition*, 2016. <https://arxiv.org/abs/1602.07261>.
- [16] CHEN L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [17] 刘致驿, 孙韶媛, 任正云, 等. 基于改进 DeepLabv3+的无人车夜间红外图像语义分割[J]. 应用光学, 2020, 41(1): 180-185.
- LIU Zhiyi, SUN Shaoyuan, REN Zhengyun, et al. Semantic segmentation of nocturnal infrared images of unmanned vehicles based on improved DeepLabv3+[J]. *Journal of Applied Optics*, 2020, 41(1): 180-185.
- [18] 于天河, 赵树梅, 兰朝凤. 结合视觉特性的红外图像增强方法[J]. 激光与红外, 2020, 50(1): 124-128.
- YU Tianhe, ZHAO Shumei, LAN Chaofeng. Infrared image enhancement method combining visual characteristics[J]. *Laser & Infrared*, 2020, 50(1): 124-128.
- [19] Zuiderveld Karel. Contrast Limited Adaptive Histogram Equalization[J]. *Graphic Gems IV. San Diego: Academic Press Professional*, 1994: 474-485. DOI: 10.1016/B978-0-12-336156-1.50061-6.
- [20] WONG S C, Gatt A, Stamatescu V, et al. Understanding data augmentation for classification: when to warp?[C/OL]//*International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2016. <https://arxiv.org/pdf/1609.08764.pdf>.
- [21] Csurka G, Larlus D, Perronnin F. What is a good evaluation measure for semantic segmentation?[C/OL]//*BMVC*, 2013. <http://www.bmva.org/bmvc/2013/Papers/paper0032/abstract0032.pdf>.