

## 基于改进的 Deeplabv3+ 的红外航拍图像架空导线识别算法

李昭慧, 寇鸽子\*

(西安科技大学 通信学院, 陕西 西安 710600)

**摘要:** 随着国家电网规模的不断扩大, 架空导线作为电力系统的重要组成部分, 对它的定期巡检变得极其重要, 同时, 随着低空飞行领域的开放, 为了保证国家电网的正常运行及低空飞行的安全, 架空导线的识别也变得极其重要。文中提出了一种使用 Deeplabv3+语义分割网络模型对红外航拍图像架空导线进行识别的方法, 并且针对红外架空导线图像目标的特征对该算法进行了改进。首先在原 Deeplabv3+算法的特征提取主干网络 ResNet50 中加入注意力机制, 使网络突出导线目标所在区域的特征, 更加关注导线目标所在的位置, 进而弱化背景等非主要区域的特征。然后对 Deeplabv3+的编码器部分进行改进, 在 ResNet50 模型中加入特征金字塔网络, 可以将浅层和深层的特征进行融合, 增强网络对不同大小目标属性的识别能力, 及导线这种小目标的检测能力, 进而提高网络的整体识别效果。实验结果表明: 改进后的算法检测性能良好, 均像素精度为 93.52%, 平均交并比为 87.83%。

**关键词:** Deeplabv3+; 特征金字塔; 架空导线; 注意力机制; ResNet50

**中图分类号:** TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20220112

## Infrared aerial image overhead wire identification algorithm based on improved Deeplabv3+

Li Zhaohui, Kou Gezi\*

(School of Communication, Xi'an University of Science and Technology, Xi'an 710600, China)

**Abstract:** With the continuous expansion of the scale of the national grid, overhead conductors, as an important component of the power system, have become extremely important to its regular inspection. At the same time, with the opening of the low-altitude flight field, the identification of overhead conductors has also become extremely important to ensure the normal operation of the national grid and the safety of low-altitude flight. The method for identifying overhead wires in infrared aerial images was proposed using the Deeplabv3+ semantic segmentation network model, and the algorithm was improved according to the characteristics of infrared overhead wire image targets in this paper. Firstly, an attention mechanism was added to the feature extraction backbone network ResNet50 of the original Deeplabv3+ algorithm, so that the network highlights the characteristics of the area where the wire target was located, and paid more attention to the location of the wire target, thereby weakening the background and other non-main area features; Then, the encoder part of Deeplabv3+ was improved, and the Feature Pyramid Networks (FPN) was added to the ResNet50 model, which can fuse the shallow and deep features, enhance the network's ability to identify the attributes of targets of different sizes, and the performance of small targets such as wires, and then improve the overall recognition effect

收稿日期: 2022-04-10; 修订日期: 2022-05-25

基金项目: 国家自然科学基金 (61705178)

作者简介: 李昭慧, 女, 讲师, 硕士生导师, 主要研究方向为光衍射特性计算与反演、图像处理。

通讯作者: 寇鸽子, 女, 硕士生, 主要研究方向为图像处理。

of the network. The experimental results show that the improved algorithm has good detection performance, the average pixel accuracy is 93.52%, and the average intersection ratio is 87.83%.

**Key words:** Deeplabv3+; feature pyramid; overhead wire; attention mechanism; ResNet50

## 0 引言

随着特高压以及高压输电等技术的飞速发展, 国家电网的规模也越来越大。架空导线作为国家电力系统的重要组成部分, 需要定期巡检, 以保证电力系统的正常运行<sup>[1]</sup>。同时, 低空领域的开放, 使得低空飞行安全成为人们关注的一个问题<sup>[2]</sup>, 而架空导线因其体积小, 在复杂环境下, 肉眼难以识别, 因此成为低空飞行过程中易引起事故的因素之一。为了保证国家电网的正常运行, 以及低空飞行领域的安全, 导线目标的识别变得极其重要。近年来, 随着无人机技术的飞速发展, 基于无人机航拍图像的架空导线的提取成为了导线检测的研究热点<sup>[3-4]</sup>。

近年来, 国内外专家对架空导线的提取做了很多的研究, 一些采用传统方法对架空导线图像进行处理和识别, 主要是通过过滤掉背景复杂环境、运动模糊等因素的影响, 通过 Hough 变换、Canny 算子、Gabor 算子等算法, 增强线型目标特征, 来实现目标分割, 完成导线目标的提取。如参考文献 [5] 采用 Hough 变换来检测传输线图像中的曲线数, 通过改变阈值, 能够有效地识别视频图像中的导线曲线进行检测。参考文献 [6] 设计了一种滤波模块, 该模板依据导线无人机航拍图像中纵向和横向的局部灰度进行设计, 并将方向滤波的结果进行自相关增强。这种方法方能够在增强架空导线目标的同时, 削弱复杂背景的干扰, 提高了架空导线的识别率, 但该算法依赖于人工控制迭代次数。参考文献 [7] 对 Canny 算子进行改进, 使用高斯滤波器和双边滤波器融合 Canny 算子结果的混合方法, 对输电导线的覆冰情况进行检测, 提高了检测的精度。可以看出传统的图像处理算法, 实时性差, 效率低下。

近年来, 深度学习在图像识别及分类等方面取得了很不错的成果。基于深度学习的图像识别算法可以分为基于目标检测的算法和基于语义分割的算法, 基于目标检测的算法适用于对具有特定区域的目标进行检测, 而导线目标细长, 一般跨越整张图片, 这导致标注的真值框充满整个图像, 会影响深度学习网络

模型获取有效区域, 因此, 目标检测并不适用于导线目标提取。而语义分割算法能够对目标实现像素级的分割<sup>[8]</sup>, 以深度卷积神经网络 (Deep Convolutional Neural Networks, DCNN) 为代表的算法, 在图像的分类<sup>[9]</sup>、图像识别<sup>[10]</sup>等方面得到了广泛应用, 该方法已用于对电力设备的识别和检测, 如绝缘子<sup>[11]</sup>、导线等<sup>[12]</sup>, 目前语义分割算法多于如遥感图像分割<sup>[13]</sup>、医学影像分析<sup>[14]</sup>、无人机控制、自动驾驶<sup>[15]</sup>、面部分割等方面, 也适用于导线目标的提取<sup>[16-18]</sup>。目前最新的 Deeplab 系列<sup>[19]</sup>语义分割算法, 将 DCNN 与条件随机场 (Conditional random field, CRF) 相结合<sup>[20]</sup>, 使得语义分割的精度更高。随后, 谷歌公司提出了 Deeplabv3+ 语义分割模型。该算法结合了之前系列算法的优点, 并引入编码器-解码器结构, 是目前分割效果最好的模型<sup>[21]</sup>。

目前, 针对架空导线识别的研究大多基于可见光图像, 由于导线具有特定的热辐射特性, 使其在红外波段显著区别于其他地物, 表现在红外图像中, 导线目标和景物存在着显著的温差和外形差异, 故文中选用红外航拍图像的架空导线图像进行识别, 提出了一种基于 Deeplabv3+网络模型的红外架空导线目标识别方法, 并根据导线目标的特征, 在原 Deeplabv3+算法的基础上进行改进, 在该算法的骨干特征提取网络 Resnet50 中加入注意力机制, 使网络有目标地进行学习, 提高导线目标识别的精准度, 同时将网络与特征金字塔网络 (Feature Pyramid Networks, FPN) 进行融合, 针对解码端在从低级特征恢复到高级特征时, 会造成细节信息丢失的问题, 增强网络的特征融合性, 提高导线目标的分割准确度。

## 1 Deeplabv3+算法

Deeplabv3+是目前最先进的语义分割算法, 文中选用该模型对架空导线的红外图像进行识别, 其网络结构图如图 1 所示, 该模型将空间金字塔池化 (ASPP) 模块<sup>[22]</sup>和编码器-解码器结构<sup>[23]</sup>相结合, 文中选用 ResNet50 作为该网络进行特征提取的主干网络, 该网络由 5 个卷积块组成 (Conv 1-Conv5\_x), 每个

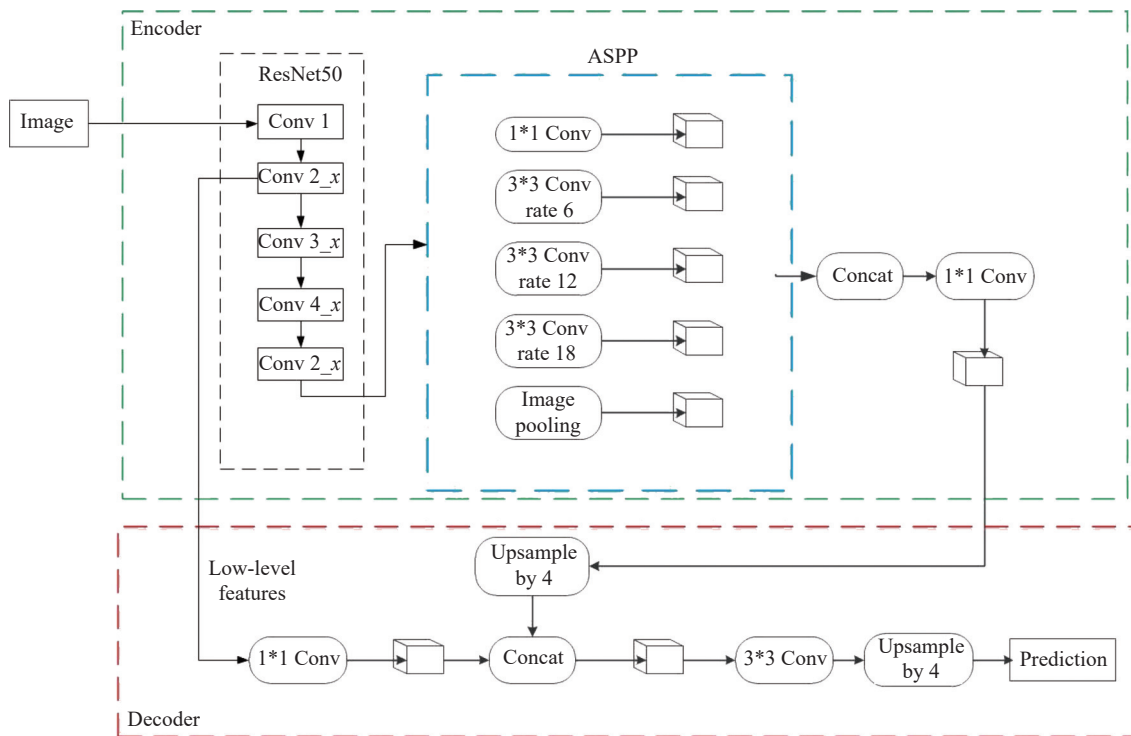


图 1 Deeplabv3+网络结构图

Fig.1 Structure diagram of Deeplabv3+ network

卷积块是由不同数量的瓶颈残差块组成。

Deeplabv3+模型采用编码器-解码器结构实现导线目标的分割,编码器主要是利用卷积神经网络对导线目标进行特征提取,输入到解码器;解码器则使用插值上采样,对编码器提取到的小尺寸特征图进行放大,恢复到图像的原尺寸,完成导线目标的分割,相比于仅有编码器的模型,编码器-解码器结构能够很好地对图像的低层次特征和高层次语义信息进行融合,很好地避免因上采样操作导致的目标像素丢失问题,能够更好的对导线目标进行分割。

编码器模块首先通过 Resnet50 网络模型对输入的红外导线图像进行特征提取,ASPP 模块通过多个不同采样率的空洞卷积并联,使网络实现从多个尺度捕获导线目标的上下文信息,最后再通过一个 1\*1 的卷积核,该卷积核的目的是为了降低通道数,得到了导线目标的高层次特征,并将高层次特征图输入到解码器中进行处理,此时的特征图尺寸变原图的 1/16。

解码器模块从基础网络 ResNet50 的第 2 个卷积块 Conv 2 中提取低层次语义特征图,对该特征图进行 1\*1 的卷积操作,减少其通道数;再对从编码器模块输出的导线目标的多尺度高层特征图进行 4 倍双

线性插值形式的上采样操作,将图像的尺寸变为输入解码器时特征图的 4 倍;然后将该特征图与经过卷积的低层次特征进行融合。将二者连接的结果再通过 1 个 3\*3 的卷积核,对融合后的特征进行特征合并,此时,图像的尺寸为原图的 1/4;最后对合并的特征图进行一次 4 倍双线性插值形式的上采样操作,图像恢复原尺寸,完成导线目标的分割。

## 2 改进的 Deeplabv3+算法

### 2.1 注意力机制

架空导线因拍摄角度的不同,导线的形状、大小以及在图像中的具体位置都是随机变化的,导线目标在整幅图像中占据很少的比例,导致背景区域在训练中占据主要地位,这使得模型的质量下降,为了解决这一问题,文本使用 Triplet Attention(TA)注意力机制<sup>[24]</sup>对 ResNet 网络模型进行改进,使网络能够自主学习,对有效的特征通道的权重进行提高,进而使网络对重要通道进行重点关注<sup>[25]</sup>。

传统的通道注意力机制在获取全局感受野的同时,会造成大量的空间信息的丢失,且通道注意和空间注意计算的分开会带来很大的计算量<sup>[26]</sup>。而 TA 与

传统通道注意力机制相比,能够很好地保留空间信息和通道信息;能够对不同维度的语义关系进行提取,实现跨纬度交互,无需进行通道和权重的间接对应;TA 几乎不引入多余参数,计算量小,能够以最小的开销提升精度<sup>[27]</sup>。该注意力机制的原理图如图 2 所示,通过 3 个分支实现跨纬相互作用,TA 通过旋转,分别捕获 3 个张量 (C,H)、(C,W)、(H,W) 之间的关系,

获取不同维度的交互信息,计算注意力的权重,第 1 个分支用来构建空间维度 H 和通道维度 C 之间的关系;第 2 个分支用来构建空间维度 W 和通道维度 C 之间的关系;第 3 个分支用来构建空间维度 H 和空间维度 W 之间的关系 (H,W),最后对 3 个分支的输出进行平均后汇总,得到精炼后的注意力张量。

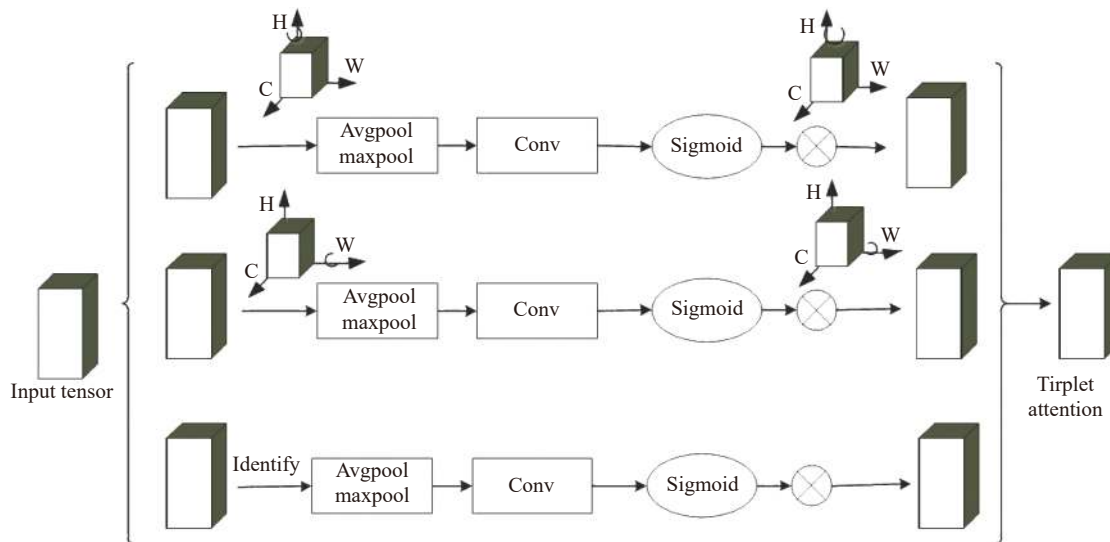


图 2 Triplet Attention 原理图

Fig.2 Schematic of Triplet Attention

笔者选择在 ResNet50 网络每个卷积块中残差模块的 shortcut 之后加入注意力机制, ResNet50 的残差模块由 2 个 1\*1、1 个 3\*3 的卷积组成,图 3 为残差网络改进的对比图,图 (a) 为原 ResNet50 的残差模块,加入 Triplet Attention 后的新的残差模块见图 (b)。

原 ResNet50 网络共包含 5 个卷积块,每个卷积块中包含不同个数的残差模块。ResNet50 网络包含 50 个 Conv 操作,能够很好地提取网络的特征,但随着网络的加深,导线目标的局部特征也会丢失,引入注意力机制后,能够使网络突出导线目标所在区域的特征,而弱化背景等非主要区域的特征,进而更好地提取导线目标。引入注意力机制的 Resnet50 的网络结构 ResNet-TA 的示意图如图 4 所示,图中输入的导线图像大小为 128×128,通道数为 3。

笔者将残差网络提取到的导线目标的特征图 (C×H×W),其中 C 为特征图的通道数, H 表示特征图的高度, W 表示特征图的宽度,以 Conv 2\_x 卷积块输

出的特征图为例,其通道数 C 为 256,宽度 W 为 64,高度 H 为 64,将这个特征图分别输入到注意力机制的 3 个分支中去。

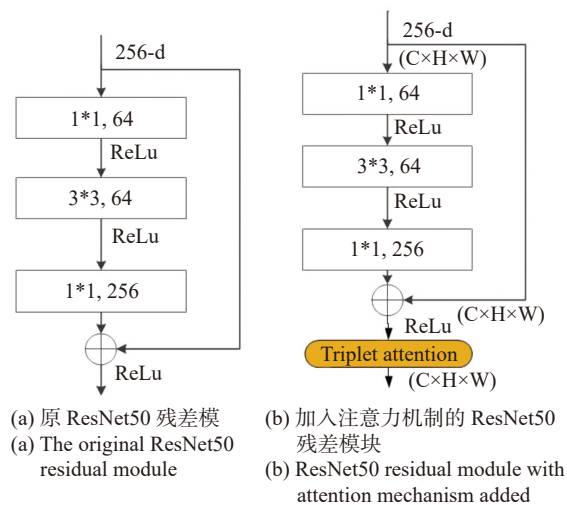


图 3 残差网络改进对比图

Fig.3 Residual network improvement comparison chart

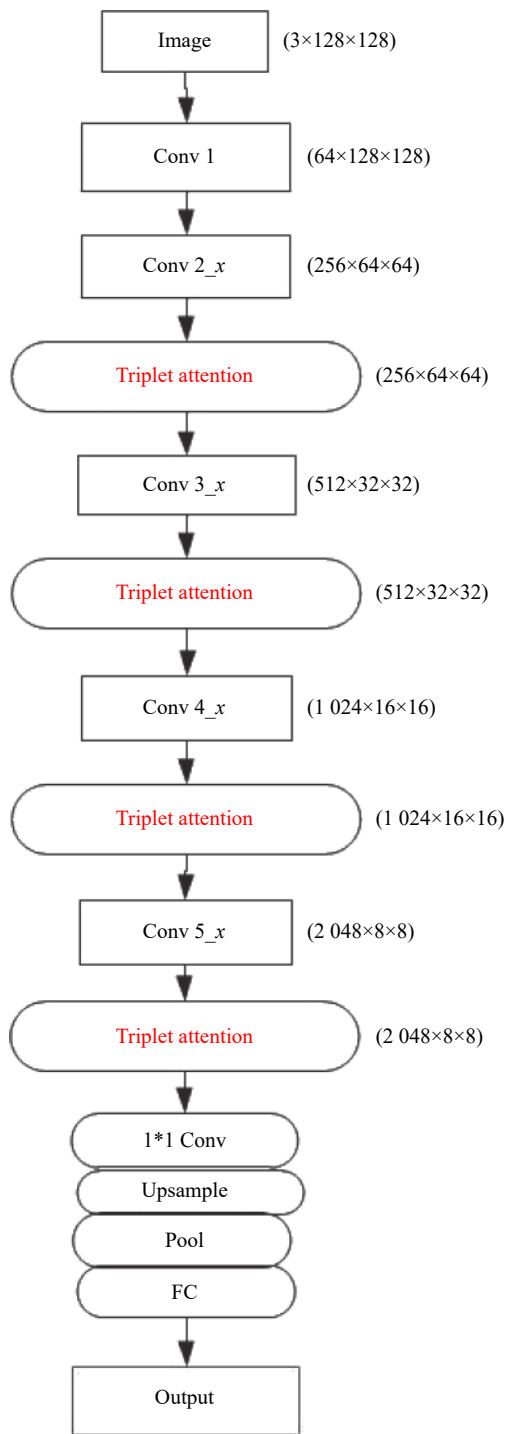


图 4 ResNet-TA 网络结构示意图

Fig.4 Structure diagram of ResNet-TA network

将得到的更有效的特征权重和原始输入的导线目标的特征图相乘,就实现了在空间维度 W 上的对原始特征的重新标定。原始网络输出的特征图对所有通道都公平对待,不同的通道描述了不同的特征,但并不是所有的通道对检测的导线目标都有意义,而

经过注意力在空间维度 W 对特征的重新标定后,使得导线所在的通道权重更高,网络能够关注导线区域所在的特征通道,进而抑制背景区域所在的非重点通道,能够使网络提取到更多的导线特征。第 2 个分支与第 1 个分支类似,用于建立通道 C 和空间 W 维度之间的交互,实现了对原始特征的重新标定。

第 3 个分支用于构建空间注意力。首先对导线特征图在通道维度上进行全局最大池化、全局平均池化操作,得到 2 个不同的特征描述,并进行合并,再经过卷积及 Sigmoid 函数,就可以得到注意力权重,将该权重与原始输入的导线特征图相乘,就可以对特征图实现在空间维度的重新标定,增强了导线目标所在的位置信息,而对其他无用的位置信息进行了抑制,使得网络更加关注导线所在的位置信息。

最后,把 3 个分支所得的重新标定的特征张量聚合再平均,实现特征在通道和空间维度的联合,使得图像中导线目标的特征更加突出,背景干扰因素得到抑制,同时也能够使网络更容易捕获导线目标所在的位置,提高网络的检测效率。

2.2 特征金字塔网络

导线目标的检测一直是一个具有挑战性的问题,这是因为图像中导线元素一般只占很少的像素,而背景像素则占据了很大的比例,且导线会以各种角度出现在任何地方。Resnet50 网络虽然能够通过卷积进行特征提取,但原始图像经过一定倍数的下采样操作后,一些小像素物体会消失在最后一层特征图中,使得后续的检测网络无法检出该目标。原 Deepbabv3+ 算法采用的融合低层特征的方法在前两次下采样时会丢失较多边界信息,并且由于双线性插值上采样的方法,图片被平滑放大,模型欠缺对目标分割的边界恢复能力。直接用于检测导线图像时,会丢失部分导线信息,极易造成漏检误检。故文中引入特征金字塔来解决这一问题。

FPN 是 Lin 等提出的<sup>[28]</sup>,该结构的提出主要是为了解决网络因多次卷积和池化操作引起的目标信息减少,高层次的特征图会丢失小目标的问题。该方法利用网络在传播时的尺度变化来建立特征金字塔,通过从上而下及横向连接的方式将得到的不同尺度的特征图进行融合,也就是将处于上层的语义信息强的低分辨率图像和处于下层的语义信息弱的高分辨率

的图像特征进行融合,进而将导线图像的语义与空间信息结合,减少像素点少的导线目标的丢失。

文中选用 ResNet50 作为主干网络,在该网络中引入特征金字塔结构, ResNet-FPN 的结构如图 5 所示,使用 ResNet50 来构建自底而上的路径, ResNet50 由 5 个卷积块构成,通过自底而上的方式进行特征提取,5 个卷积块 (Conv 1-Conv 5\_x) 输出的特征图路径为 {C1,C2,C3,C4,C5}, 这个过程中,图像的空间维度每次减为原来的二分之一,连接相邻特征图的通道加倍。自顶而下的路径是将顶层的特征图采取自顶向下的方式进行反卷积,将经过 2 倍上采样后的特征图与相邻层进行逐像素加合,进行特征融合。这个过程中,首先对 C5 进行 1\*1 的卷积操作得到特征图 M5,其通道深度降为 256,然后按照自顶而下的路径进行,图像的空间维度会因为 2 倍上采样扩大两倍,同样的 C4、C3、C2 也要经过 1\*1 的卷积,使得它们的通道数也变为 256,依次得到 M4、M3、M2 特征图,然后对 M5、M4、M3、M2 进行 3\*3 的卷积操作,来消除因逐像素加合引起的混叠效应,得到最后的特征图 P2、P3、P4、P5,这些特征图的通道数都为 256,但是尺寸不同。这样 FPN 就是融合了底层到高层的特征图,

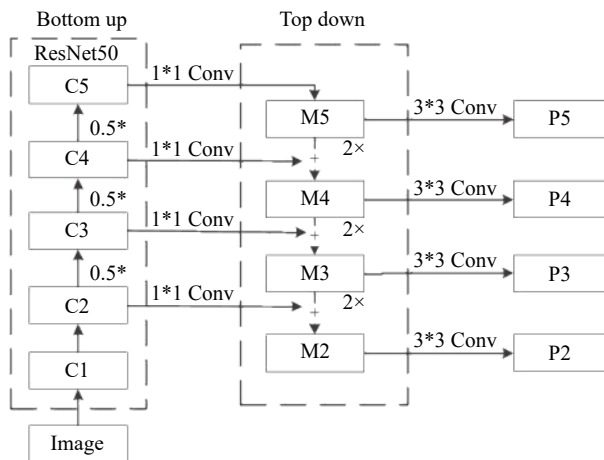


图 5 ResNet-FPN 结构图

Fig.5 Structure diagram of ResNet-FPN

进而提取到了各个阶段的特征,提高了导线这种小像素点目标的检测能力。

### 2.3 整体模型

结合特征金字塔和 Triplet attention 的优点并融合特征金字塔结构,提出了文中的模型,如图 6 所示。原 Resnet50 共有 16 个残差块,每个残差块由 2 个 1\*1 和 1 个 3\*3 的卷积组成,文中在其残差模块中插入注意力机制 (Triplet attention),构成新的残差模块,

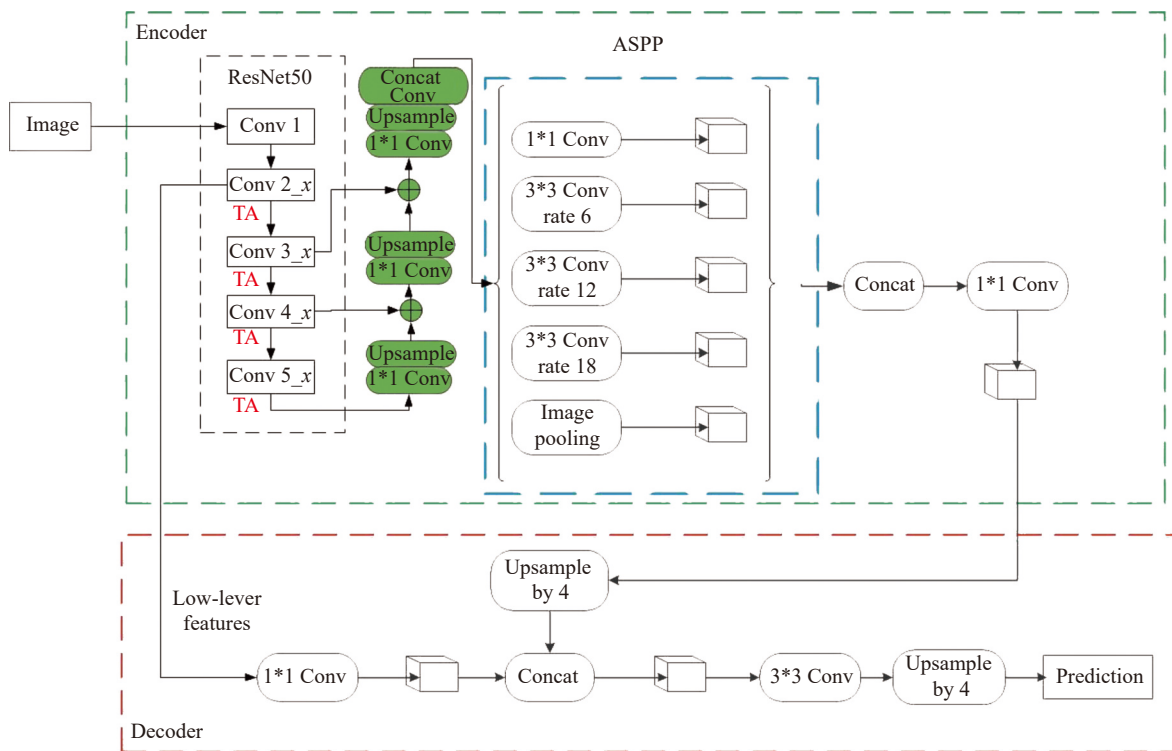


图 6 改进的 Deeplabv3+结构图

Fig.6 Structure diagram of improved Deeplabv3+

然后在 Deeplabv3+ 的主干网络 Resnet50 加入特征金字塔 (FPN), 在原本的网络结构中, 仅提取 Conv 5 块中的特征图到 ASPP 模块中, 而改进后的网络, 通过特征金字塔结构提取到 Conv 3、Conv 4、Conv 5 块的多尺度的特征图输入到 ASPP 模块中, 对特征进行融合的同时提高像素点少的小目标导线的检测效果。

### 3 实验与分析

#### 3.1 实验环境

实验硬件和软件环境如表 1 所示。

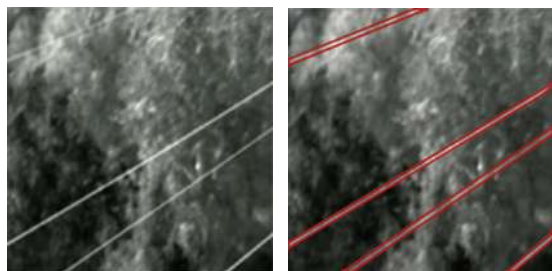
表 1 实验平台参数设置

Tab.1 Experimental platform parameter settings

Experiment platform	Model parameters
Operating system	Ubuntu18.4
CPU	I9-12900 KF 3.2 GHz
GPU	RTX2070 S 8 GB
Graphics card	12 G
Frame	Pytorch
Programming language	Pyhton

#### 3.2 数据集建立

由于目前并没有导线的公开数据集, 故文中建立了专用于红外图像架空导线检测的标签数据集。文中利用双光无人机, 采集 1000 张红外架空导线图片, 作为数据集的来源。图像尺寸统一修改为 128×128, 其中, 将图像中的导线标记为 wire, 所有图片标签均通过 Labelme 软件自行标注, 生成 json 文件, 并用于训练, 图 7 为训练样本的原始图像及标签图像。按照 8 : 2 的比例对数据集和测试集进行划分, 即 800 张用于训练, 200 张用于测试。



(a) 原始图像 (a) The original image (b) 标签图像 (b) Label image

图 7 样本原始图像及标签图像

Fig.7 Sample original image and label image

#### 3.3 数据分析

##### 3.3.1 评判标准

为了验证文中构建的导线分割模型的性能, 需选取合适的性能评价指标, 在语义分割领域有多种评价精度的指标, 文中选取均像素精度 (Mean Pixel Accuracy, MPA)、平均交并比 (Mean Intersection over Union, MIoU) 这 3 个指标作为衡量模型分割精度的标准。其中 MPA 表示所有类别像素精度的均值, 是语义分割领域中最普遍的指标之一, 如公式 (1) 所示。MIoU 表示的是真实值与预测值的交集比并集, 如公式 (2) 所示。其中 MPA 侧重表现像素级的准确率, 而 MIoU 更加侧重于表现模型计算分割区域的完整性以及分割位置的准确性。

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (1)$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{i=0}^k P_{ij} + \sum_{i=0}^k P_{ji} - P_{ii}} \quad (2)$$

式中:  $k$  表示类别总数;  $p_{ij}$  表示实际像素类别  $i$  被预测为类别  $j$  的像素数量;  $p_{ii}$  表示实际像素类别为  $i$ , 被预测为类别  $i$  的像素数量;  $k+1$  表示分割类别的数量, 其中含有一个背景类。

##### 3.3.2 算法效果评估

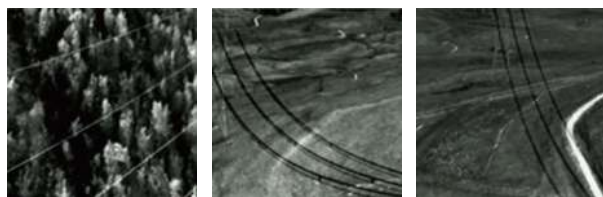
为说明该算法的有效性, 文中先使用传统算法 (Canny 算子) 对导线目标进行检测, 进行实验对比, 对 30 幅导线图的导线条数做了统计, 人眼可看到的导线一共有 132 根, 利用 Canny 算子一共检测到 108 根导线, 检测率是 81.8%, Canny 算子能够检测到部分边缘, 但是图像中还存在大量噪声, 且会丢失部分导线信息, 检测效果不佳。而后使用文中提出的算法, 实验表明, 该算法具有良好的检测效果。文中先使用原 Deeplabv3+ 的网络结构对导线目标进行检测, 然后对 Deeplabv3+ 的特征提取主干网络 ResNet50 进行改进, 引入注意力机制, 再引入特征金字塔网络, 再次进行实验, 实验结果如表 2 所示。

3 种检测算法的实验结果对比如图 8 所示, 其中图 (a) 为原图; 图 (b) 为 Canny 算子检测结果, 从图中可以看出, Canny 算子的检测存在大量噪声, 且导线信

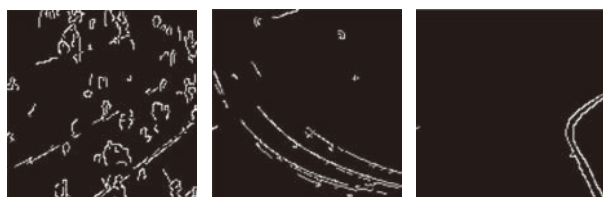
表 2 实验结果度量

Tab.2 Experiment results metrics

Method	MPA	MIOU
Deeplabv3+	92.91%	86.79%
Proposed algorithm	93.52%	87.83%



(a) 原图  
(a) Original image



(b) Canny算子检测结果图  
(b) Canny operator detection result graph



(c) 改进前检测结果图  
(c) Test results before improvement



(d) 改进后检测结果图  
(d) Improved post-test results

图 8 实验结果对比图

Fig.8 Comparison chart of experimental results

息提取不完整,在道路这种具有热辐射的图像中,该算子会将道路误检为导线目标;图(c)为原 Deeplabv3+可以较好地提取到导线信息,但对于特别细小的导线目标提取不完整,这是因为原 Deeplabv3+算法在进行特征提取时,随着网络的加深,一些细小的导线目标会消失在最后一层特征图中,导致后续的网络无法检测到该目标;图(d)为文中的改进算法,结合了注意力

机制及特征金字塔,改善了特征提取过程中,像素点少的细小的导线目标丢失的问题,能够完整地提取到导线信息,效果最好。

## 4 结论

文中针对红外图像架空导线的检测问题,提出了基于 Deeplabv3+的改进算法,使用 Labelme 对数据集进行标注。对 Deeplabv3+编码器的特征提取主干网络 ResNet50 进行改进,引入注意力机制,增强了网络的目标特征。并针对解码端在从低级特征恢复到高级特征时,会造成细节信息丢失的问题,引入特征金字塔网络,增强网络的特征融合性,提高导线目标的分割准确度。实验表明,文中提出的算法均像素精度为 93.52%,平均交并比为 87.83%,能够精准地实现对导线目标检测。该算法可用于对无人机巡检的图片实现快速检测导线目标。文中所做研究为基于目标区域的导线区域识别,未来将基于划定目标,实现单根导线目标的识别,并继续深入研究针对无人机航拍视频中的导线目标识别,实现视频流中导线目标的实时检测和跟踪。

## 参考文献:

- [1] Du Zhongming, Wang Xuesong. Forecast of China's electricity demand level during the "13th Five-Year Plan" [J]. *China Electric Power*, 2017, 50(9): 11-17. (in Chinese)
- [2] Pouliot N, Richard P L, Montambault S. Line scout technology opens the way to robotic inspection and maintenance of high-voltage power lines [J]. *IEEE Power and Energy Technology Systems Journal*, 2015, 2(1): 1-11.
- [3] Zhang Congxin, Zhao Le, Wang Xianpei. A fast extraction algorithm of power lines under the background of complex ground objects [J]. *Journal of Wuhan University (Engineering Edition)*, 2018, 51(8): 732-739. (in Chinese)
- [4] Zhao Haocheng, Lei Junfeng, Wang Xianpei, et al. Algorithm for power line recognition in aerial images with complex background [J]. *Bulletin of Surveying and Mapping*, 2019(7): 28-32. (in Chinese)
- [5] Han C, Liu Q. Appearance monitoring of the transmission lines based on hough transform[C]//2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2020: 71-75.
- [6] Weiran C, Zhu Linlin, Jianda H. Iterable multidirectional



- autocorrelation approach for aerial power line image enhancement [J]. *Robot*, 2015, 37: 738-747.
- [7] Nusantika N R, Hu X G, Jin X. Improvement Canny edge detection for the UAV icing monitoring of transmission line icing[C]//2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), 2021: 1838-1843.
- [8] Shelhamer E, Jonathan L, Trevor D. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [9] Vo A T, Tran H S, Le T H. Advertisement image classification using convolutional neural network[C]//2017 9th International Conference on Knowledge and Systems Engineering (KSE), 2017: 197-202.
- [10] Ding J, Chen B, Liu H, et al. Convolutional neural network with data augmentation for SAR target recognition [J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(3): 364-368.
- [11] Luo Jianjun, Liu Zhensheng, Gong Xiang, et al. State evaluation method of line insulators based on UAV images and transfer learning [J]. *Electric Power Engineering Technology*, 2019, 38(5): 30-36. (in Chinese)
- [12] Madaan R, Maturana D, Scherer S. Wire detection using synthetic data and dilated convolutional networks for unmanned aerial vehicles[C]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017: 3487-3494.
- [13] Yuan Li, Yuan Jishou, Zhang Dezheng. Remote sensing image classification based on DeepLab-v3+ [J]. *Advances in Lasers and Optoelectronics*, 2019, 56(15): 152801. (in Chinese)
- [14] Ha D, Riegler M A, Johansen D, et al. Doubleu-net: A deep convolutional neural network for medical image segmentation [C]//2020 IEEE 33rd International Symposium on Computer Based Medical Systems (CBMS), 2020: 558-564.
- [15] Sun L, Yang K, Hu X, et al. Real-time fusion net-work for RGB-D semantic segmentation incorporating unexpected obstacle detection for road-driving images [J]. *IEEE Robotics and Automation Letters*, 2020, 5(4): 5558-5565.
- [16] Zhao Le, Wang Xianpei, Yao Hongtai, et al. A review of power line extraction algorithms based on visible light aerial images [J]. *Power System Technology*, 2021, 45(4): 1536-1546. (in Chinese)
- [17] Zheng Kai, Li Jiansheng. A review of image semantic segmentation based on deep neural network [J]. *Surveying and Mapping and Spatial Geographic Information*, 2020, 43(10): 119-125. (in Chinese)
- [18] Xu Hui, Zhu Yuhua, Zhen Tong, et al. A review of deep neural network image semantic segmentation methods [J]. *Computer Science and Exploration*, 2021, 15(1): 47-59. (in Chinese)
- [19] 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.
- [20] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[EB/OL]. (2014-12-22)[2022-04-10]. <https://arxiv.org/abs/1412.7062>.
- [21] Wang Yagang, Xi Yiyuan, Pan Xiaoying. Improved deep Labv3+ network for intestinal polyp segmentation [J]. *Computer Science and Exploration*, 2020, 14(7): 1243-1250. (in Chinese)
- [22] 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*, 2017, 40(4): 834-848.
- [23] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[EB/OL]. (2014-09-10)[2022-04-10]. <https://arxiv.org/abs/1409.3215>.
- [24] Misra D, Nalamada T, Arasanipalai A U, et al. Rotate to attend: Convolutional triplet attention module[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3139-3148.
- [25] Wang F, Tax D M J. Survey on the attention based RNN model and its applications in computer vision[EB/OL]. (2016-01-25)[2022-04-10]. <https://arxiv.org/abs/1601.06823v1>.
- [26] Hou Ruihuan, Yang Xiwang, Wang Zhichao, et al. A real-time detection method for forestry pests based on YOLOv4-TIA [J]. *Computer Engineering*, 2022, 48(4): 255-261. (in Chinese)
- [27] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2020: 10781-10790.
- [28] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 2117-2125.