

一种面向复杂场景的图像边缘检测方法

李金中¹, 左宇晨², 汪玉¹, 王子磊²

¹国网安徽省电力有限公司电力科学研究院, 安徽 合肥

²中国科学技术大学, 安徽 合肥

收稿日期: 2021年10月1日; 录用日期: 2021年10月29日; 发布日期: 2021年11月5日

摘要

针对较为复杂的光伏电站场景中的图像中光伏板边缘提取问题, 本文提出了一种精细检测方法。所提出的方法首先引入多尺度的图像特征, 通过鼓励网络不同层输出的特征表示检测对应该层尺度的图像边缘, 进而充分融合不同尺度特征中所包含的丰富信息, 使得对于不同尺度物体边缘的检测都能够更加精细; 其次, 本方法利用多任务学习结构, 通过挖掘语义分割和边缘检测两个任务之间的相关性与互补性, 使得图像特征更加聚焦于待提取边缘的物体, 排除场景中无关物体或噪声的干扰, 从而生成更加合理的图像边缘。量化的实验结果与可视化结果均表明, 本方法能够获得更加精准的图像边缘检测效果。

关键词

多任务学习, 语义分割, 边缘检测, 计算机视觉

An Image Edge Detection Method for Complicated Scenes

Jinzhong Li¹, Yuchen Zuo², Yu Wang¹, Zilei Wang²

¹State Grid Anhui Electric Power Co., Ltd. Electric Power Research Institute, Hefei Anhui

²University of Science and Technology of China, Hefei Anhui

Received: Oct. 1st, 2021; accepted: Oct. 29th, 2021; published: Nov. 5th, 2021

Abstract

Aimed at edge extraction of photovoltaic panels in the complicated photovoltaic power stations, this paper puts forward a precise detection method. For the proposed method, first, multi-dimensional image features are introduced. The feature representations from different layers of network detect image edge correspond to the dimension of this layer. The plenty of information in different dimension features is fully integrated to make the edge detection of object with different

dimensions be more precise. Second, this method makes use of multi-task learning structure and digs out the correlation and complementarity between semantic segmentation and edge detection to make low-level features of image more concentrate on the object edge to be extracted and eliminate the interference of irrelevant objects or noise in the scene, thereby generating a more ideal image edge. Both quantitative experimental result and visualization result show that this method can acquire more precise image edge detection effect.

Keywords

Muti-Task Learning, Semantic Segmentation, Edge Detection, Computer Vision

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

边缘检测是图像处理和计算机视觉领域中的一项基本任务，该任务的目的是标识数字图像中亮度变化明显的点的集合。在例如视觉显著性检测，医学图像处理，以及自动驾驶等在内的许多现实应用场景中，图像边缘检测方法都是作为一项十分重要的基础性能力[1] [2] [3] [4] [5]。

传统的边缘检测方法如 Canny [6]、Sobel [7]等首先使用预定义好的特征描述子提取图像中相邻像素间差异较大的位置，而后对结果进行滤波等后处理操作。但是，由于预先定义好的特征描述子不能够很好地适应多样的视觉图像，检测结果往往容易出现边缘不够连续的问题，并且这些方法的效果依赖于参数的人工调试，设定不同阈值所得到的边缘效果往往天差地别。此外，图像中的一个物体内部也有可能存在亮度变化较为显著的位置，传统方法容易在这些位置处也提取出图像边缘，而不会根据语义信息自适应地调整所得到的结果。随着深度卷积网络在计算机视觉领域取得成功，研究人员逐渐转向使用神经网络作为图像的边缘提取器。针对传统方法检测出的边缘不够连续等问题，之前的方法提出了 HED [8]、BDCN [9]等边缘检测模型，在一些较为简单的场景下取得了一定的效果。但一方面，当待提取边缘的物体尺寸较小或边缘较为精细时，由于神经网络的骨干特征提取器对于输入图像进行下采样，因此这些模型很难直接利用所提取的特征图对小物体进行检测，导致其在精细的物体边缘边缘检测任务上往往表现不佳。另一方面，当场景中的其他物体或物体本身所带线条等干扰项过多时，神经网络难以直接根据语义自适应地判断需要提取哪些位置的边缘，进而模型得到的边缘也会包含这些干扰项。

本文首先考虑一种复杂场景下——光伏电站场景中的光伏板图像边缘检测任务。具体而言，在光伏电站中，为了实现对光伏板的智能巡检，往往首先需要在无人机的拍摄图像内自动定位出光伏板的位置，而考虑到光伏板本身形状不规则、语义分割需要更高的计算量，难以满足电站智能巡检对于时延的要求等因素，边缘检测更好地适配了在智能巡检过程中对于光伏板实时精确定位的工作。在这一任务下，我们首先根据实地拍摄的无人机巡检图像构建了光伏板边缘检测数据集，包括原始图像与人工标注的光伏板边缘；其次，我们注意到在这一场景中，无人机飞行高度较高、光伏板排布密集，因而边缘部分的微小误差都会导致对于光伏板定位的错误，需要提取更加精细的边缘图像；另外，实际的光伏电站中包含了电线杆、树木等一些其他物体，每块光伏板中间的线条也容易被模型视为光伏板的边缘，之前的方法难以排除这些不相干物体的干扰，因而难以利用边缘检测图像实现对光伏板的准

确定位。

针对以上问题, 本文提出面向复杂场景的图像边缘检测方法。一方面, 本方法使用特征提取器多层的特征输出, 构建多尺度的图像特征, 利用不同尺度的特征图检测对应尺度的光伏板, 尤其是在底层特征中重点关注图像中尺寸较小、表征不明显的光伏板, 从而提升边缘检测模型在小物体上的表现; 另一方面, 本方法使用多任务学习的思想, 在骨干网络共享语义分割与边缘检测的模型参数, 进而共享两个任务所学到的信息, 使边缘检测模型更关注两个问题之间丰富的关联特征, 提升模型对含有语义信息的物体边缘的筛选能力。综上所述, 本文所提出的创新之处在于:

a) 本文提出一种面向复杂场景的新的边缘检测任务, 该任务要求根据光伏电站的无人机航拍图像自动地提取图像中的光伏板图像边缘。相比于其他的图像边缘提取任务, 本文提出的任务更加复杂且实用: 一方面图像中的该任务需要边缘检测模型更加精准且能够排除无关因素的干扰, 另一方面该任务能够直接应用在对于光伏电站的无人机智能巡检之中。

b) 为了解决复杂场景下的图像边缘检测问题, 本文从多尺度特征与多任务学习两个角度对原有的边缘检测方法进行改进。首先, 使用骨干神经网络提取多尺度特征, 利用底层特征保存较高图像分辨率的特性来重点对检测结果进行精细调整; 而后, 利用语义分割任务所能够提供的语义信息, 使得模型对于物体边缘的提取能力更加有针对性。

c) 为了检验算法提升效果, 本文构建光伏板图像边缘数据集, 并在该数据集上验证方法效果。大量实验结果表明, 本文提出的方法在定量指标与可视化结果两方面都超过了之前的方法, 且分析实验结果表明本文所提出的两个模块分别对于最终的边缘检测结果起到了一定的提升作用。

2. 相关工作

2.1. 传统边缘检测算法

传统的边缘检测方法包括 Sobel 和 Canny 边缘检测算法等, 这些方法的核心思想是寻找待测图像的亮度梯度, 并为该梯度的大小设置阈值, 从而将图像分为高于阈值和低于阈值的二值图像。二值图像上每个点的值表示该点是否是原始图像中的边缘点, 1 表示该点是边缘点, 而 0 表示该点不是边缘点。计算图像亮度梯度的算子都是预先设置好, 与具体的输入图像无关, 而由于特征算子不能够根据输入图像自适应地调整, 因此往往需要人工调节阈值参数, 进而使模型能够适应不同的数据集。但由于阈值参数的设置本身仍然依赖于人工经验和每张输入图片的自身性质, 使得实际的效果较差。图 1 给出了一张 Canny 边缘检测方法在本文提出的光伏图像数据集的检测结果示例。可以明显看出, 由于待测图像中包含大量草地、树木等不需要被检测边缘的物体, 而传统的特征算子无法对这些物体做区分, 例如检测结果右半部分中密密麻麻的树木边缘。另外, 图中的光伏板虽然被正确的提取出边缘, 但是由于阈值的设置问题, 光伏板内部的线条也被当作边缘提取出来。综上所述, 传统的边缘检测方法对于边缘的筛选能力不够, 进而使得提取的边缘效果较差。

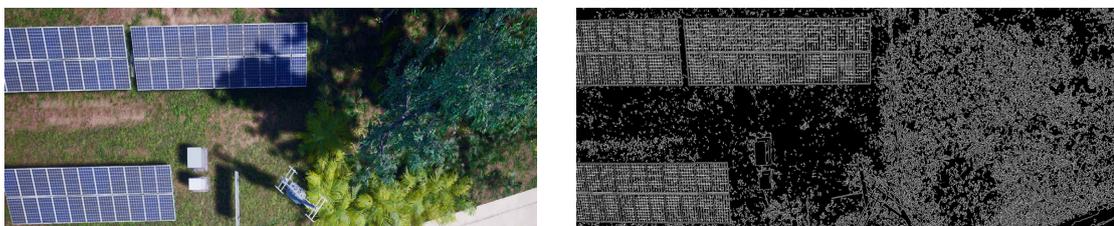


Figure 1. Illustration of Canny edge detection
图 1. Canny 边缘检测示例

2.2. 基于深度学习的边缘检测

近些年来,随着深度学习的发展,卷积神经网络模型(CNN)通过堆叠神经网络,可以自动建模图像数据中的隐藏关系,提取更具表达能力的特征,因为被越来越多的用于目标检测、语义分割等计算机视觉任务上。基于卷积神经网络的边缘检测方法的核心思想是利用神经网络拟合高斯热图,并使用热图来表示边缘点在图像上每个位置出现的概率,这种方式相当于图像上的每个位置都提供了监督信息。文献[8]提出的 HED 模型首先提出采用多尺度特征输出的网络结构,分别生成对五个不同尺度特征图学习的分支,通过权重融合函数在输入图像尺寸上得到最终的边缘输出。上述多尺度特征输出的思想既可以利用模型中的低分辨率特征图所蕴含的高维语义信息,又可以利用高分辨率特征图中的图像细节。文献[9]提出的 BDCN 模型则进一步发扬了多尺度特征输出的思想,利用双向级联网络中双向损失监督的方式,令网络中的每个模块捕捉到适合自身的边缘尺度,在压缩模型计算量的同时进一步提高了所提取边缘的准确率。

2.3. 多任务学习

文献[10]所提出的多任务学习方法,利用不同学习任务之间相关的训练信息来提升单个任务的性能。神经网络的底层参数在不同任务之间共享,进而方便利用不同任务之间的相关性来提升骨干网络的泛化能力;而在模型的顶层,各个任务优化各自任务特定的输出模块,从而使用独有参数优化各自任务的结果。多任务学习方法被提出之后,广泛地应用于包括推荐系统、计算机视觉在内的人工智能各个领域。文献[11]提出了一种新的多任务学习体系,通过学习任务特定的特征注意力,在图像分类和检测等任务上都取得了良好的性能;文献[12]设计了 Cross-Sisten 单元,利用神经网络自动判断特征图中的哪些区域应被多个任务共享,哪些区域应只应用于某一个任务上。

3. 本文方法

3.1. 问题设定

为了使边缘检测模型有针对性地学习图像上具有明确语义信息的边缘,我们对边缘的定义方式需要满足以下条件:所定义的边缘要保证对应于一个具有明确语义信息的物体,该物体同时对应一张边缘标注和语义分割图像。但是,考虑到物体边缘的判断往往存在比较大的个体差异,因此在边缘检测指标的计算时,为了避免由不同人对边缘的差异性认知所带来的偏差,本文将每张图像由多人标注后再汇总到一个标签。下图 2 给出了一组本文对于光伏电站图像的边缘和语义标注示例。

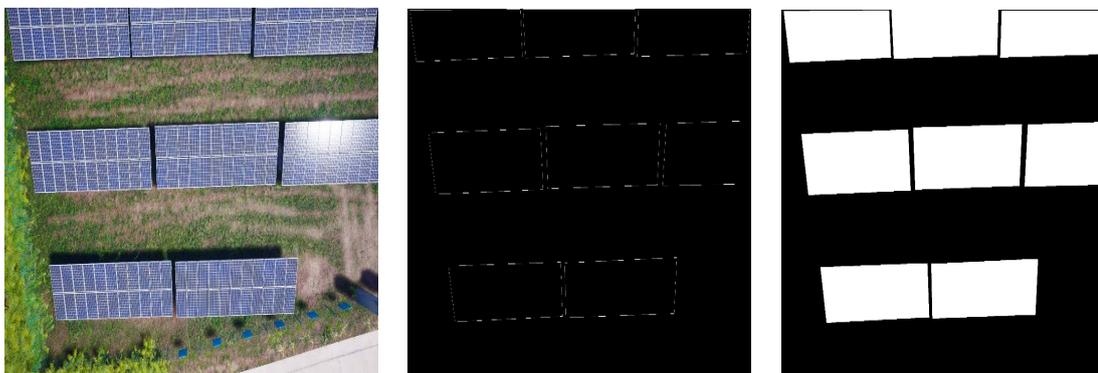


Figure 2. From left to right are the original image, edge annotation and semantic annotation correspondingly
图 2. 从左到右分别是原始图像,边缘标注和语义标注

3.2. 面向复杂场景的边缘检测

3.2.1. 整体模型

本文借鉴文献[13] [14] [15]所提出的模型, 对于每张输入图像 x , 输出 20 张高斯热图用以表达图像边缘所在的位置。其中 14 张对应于根据不同尺度特征提取得到的边缘高斯热图, 3 张对应于作为语义分割结果的高斯热图, 另外 3 张则是将语义分割热图转换为边缘高斯热图的结果。注意到, 这里所提到的高斯热图是使用数值大小来表示每个像素点和边缘或语义相关程度的矩阵, 1 与 0 分别代表相关与无关。

如图 3 所示, 对包含 M 个下采样层的骨干网络, 我们分别利用每层的特征输出, 为边缘检测任务融合不同尺度的边缘信息。为了与其他方法公平比较, 本文使用 VGG16 [16] 网络作为骨干特征提取器, 骨干网络的下采样层数量 $M = 5$ 。此外, 本文的边缘检测模型将每一层骨干网络提取的特征图分别送入语义分割和边缘检测的特征提取模块, 利用两个任务之间的相关性, 完成对边缘的准确定位和精确描绘。而为了学习到最优的边缘提取模型, 进一步引入语义转边缘(sematic-to-edge, s2e)模块。这一模块通过提取高频分量的方式将语义分割结果转换为图像边缘, 模型将 s2e 模块的输出融合至边缘检测模型的结果中, 对边缘检测的结果进行增强。下面我们具体介绍本文模型中的各个模块。

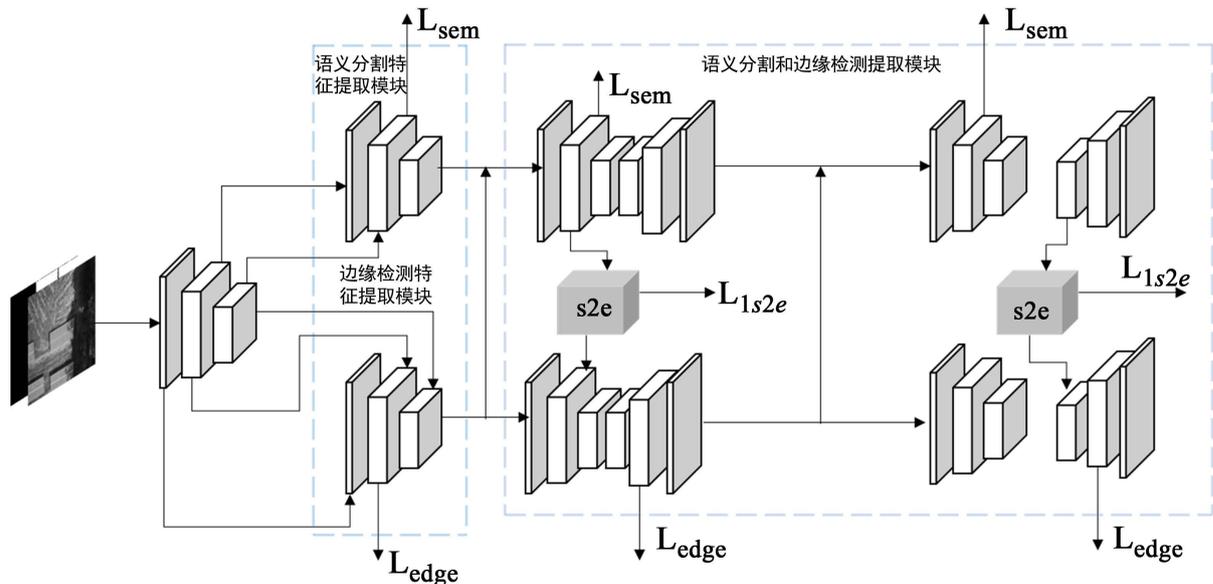


Figure 3. Feature extraction of our model

图 3. 模型特征提取部分

3.2.2. 多尺度图像特征提取

考虑到 VGG16 网络对输入图像进行了 32 倍的降采样, 因而边缘上的微小细节差距最终往往偏差巨大, 需要利用保留了图像分辨率的低维特征图来对精细的物体边缘进行识别。之前的边缘检测方法同样也考虑到了引入多尺度的特征提取, 但是它们往往使用相同的损失函数约束这 M 层的多尺度输出。相比之下, 本文希望模型从各个尺度的特征中针对性地抽取包含在该尺度特征中的判别性信息, 并对不同尺度特征的融合过程进行精细化调整, 从而进一步提升模型的边缘检测能力。具体而言, 首先针对第 i 层的特征图 F_{i1} 与 F_{i2} 生成对应的上采样图 A_{i1} 与 A_{i2} , 将特征图的尺寸恢复到与原始输入一致。其中, 对于每一个下采样层所包含的两个 $3 * 3$ 卷积, 第一个卷积输出特征作为 F_{i1} , 而将第二个卷积输出的特征作为 F_{i2} :

$$A_{ij} = g\left(\text{aspp}\left(f\left(F_{ij}\right)\right)\right) \quad (1)$$

其中， $1 * 1$ 卷积操作 $f()$ 将 F_i 的通道数量转换为指定值， aspp 模块包含几个并行的空洞卷积，对于特征图以不同采样率的空洞卷积进行采样，相当于以多个比例捕捉图像的上下文， g 代表 $1 * 1$ 卷积操作。然后，我们将各个尺度的采样图 A_{ij} 分别相加，作为输出的特征图 D_{i1}, D_{i2} 。考虑到 D_{i1}, D_{i2} 的生成过程融入了其他尺度信息，在网络的梯度反传阶段，在 D_{ij} 处将梯度进行截断。最后，将所有尺度得到的共 10 个 D_{ij} 信息进行拼接，再经过 $1 * 1$ 卷积获得边缘检测部分所输出的预测结果。

语义分割模块则是使用了基础的 Deeplab V3 [17] 架构，利用与边缘提取模块同样的骨干网络输出的特征图 F 与骨干网络中 4 倍下采样后的特征图 F_4 ，得到模型对于语义分割任务的结果输出：

$$E = \text{up}_4\left(\text{concat}\left(\text{up}_4\left(\text{aspp}\left(F\right)\right), F_4\right)\right) \quad (2)$$

其中， up_4 代表 4 倍率的双线性插值上采样， concat 代表拼接操作。

3.2.3. 结果融合模块

除了在模型的骨干特征提取网络部分共享语义分割与边缘检测模块的模型参数之外，为了更加充分地利用多任务学习的思想，本文对两模块所生成的检测结果进行融合，模型通过充分利用语义分割图对于不同类别物体的敏感性以促进对边缘的学习，同时边缘提取任务也反过来帮助语义分割模型在物体边缘处容易得到正确的分割结果。如图 4 所示，利用一个 s2e 模块来对语义分割与边缘检测模块的输出结果之间进行转换。对于语义分割的结果， s2e 模块利用全局平均池化的手段得到信号中的低频分量，进而在分割结果中减去池化结果以针对性抽取图像中的高频分量，这些高频分量对应着图像的边缘部分。而后，我们一方面将这种方式得到的边缘的检测结果与边缘检测模块直接输出的结果进行融合，二者拼接后利用 $1 * 1$ 卷积得到融合的边缘结果，注意到这里我们利用 $1 * 1$ 卷积核为不同的通道设置权重，在融合结果的输出中对于语义分割模块转换而来的与模型直接得到的边缘结果作加权平均；另一方面，为了使用语义分割任务来约束边缘检测的结果，将两个模块的直接输出拼接后送入另外的 $1 * 1$ 卷积模块，希望该模块输出两任务融合后的语义分割结果图。最终，模型的损失对于语义分割和边缘检测这两个任务的直接输出结果与融合结果分别进行约束，从而充分利用语义与边缘两个任务所包含的互补的信息。

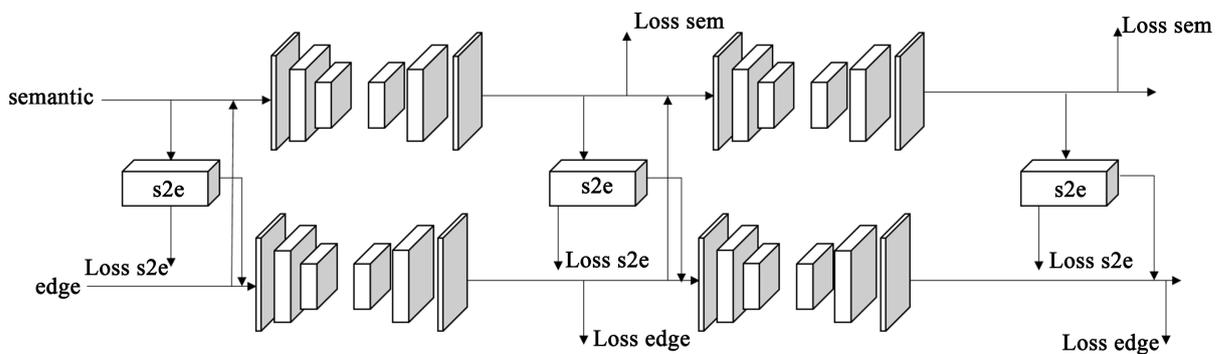


Figure 4. Result fusion module
图 4. 结果融合模块

3.3. 模型训练与测试

为了分别约束语义分割、边缘检测和 s2e 模块，对于模型的所有直接输出与融合结果输出，本文使用相同的三部分损失函数来训练模型。针对语义分割模块，计算模型预测的语义图与标签语义图之间的

加权交叉熵损失。具体地，对于出现在图像 x_i 中的类别 k ， y_i^{jk} 表示在 x_i 所对应的标签语义图上，第 j 个像素点是否对应于类别 k 的物体。即标签语义图上物体范围内的像素点被设置为 1，其余像素点置为 0。使用交叉熵损失计算图像上每个像素点的分类损失，并通过权重来减轻不同类别物体出现频率分布不均所带来的影响。损失函数 L_{sem} 为：

$$L_{sem} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{H*W} \sum_{k=1}^K w_i^k y_i^{jk} \log_e h_j^k(x_i) \quad (3)$$

$$w_i^k = \frac{H*W - P_i^k}{H*W} \quad (4)$$

其中， $h_j^k(x_i)$ 表示模型预测的语义分割结果中样本 x_i 第 j 个像素属于类别 k 的概率， K 表示场景中物体的总类别数， H 与 W 分别为输入图像的高和宽，二者相乘即为输入图像的像素数量，而 P_i^k 则表示图像 i 上标签类别为 k 的像素个数。通过这一针对不同类别损失的加权机制，在各个类别的样本数量分布不均衡的情况下，模型可以对于出现次数较少的类别获得更好的判别效果。类似地，对于边缘检测结果计算模型预测热图与标签热图之间的加权交叉熵损失。假定 E_i^k 为图像 i 上第 k 类物体边缘像素的数量，用 L_{edge} 表示这部分损失函数，如下所示：

$$L_{edge} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{H*W} \sum_{k=1}^K w_i^{k'} y_i^{jk'} \log_e h_j^{k'}(x_i) \quad (5)$$

$$w_i^{k'} = \frac{H*W - E_i^k}{H*W} \quad (6)$$

最后，我们通过使用边缘标签进行 L1 正则的方式，使得 s2e 模块的输出与真实的边缘结果保持一致。假定 $h(x_i)$ 表示语义分割的结果图像，将其减去自身平均池化后的像素值得到 s2e 模块预测的边缘。为此，计算预测的边缘结果与边缘标签热图之间的 L1 正则损失：

$$L_{s2e} = \left| y_i' - \left| h(x_i) - \text{avg_pool}(h(x_i)) \right| \right| \quad (7)$$

其中， y_i' 表示标签的边缘热图， avg_pool 表示均值池化操作，语义分割图像减去均值池化后的高频分量即为 s2e 模块输出的边缘检测结果。本文使用这三部分损失的加权求和来训练模型：

$$\text{Loss} = L_{edge} + \lambda_1 L_{sem} + \lambda_2 L_{s2e} \quad (8)$$

测试阶段，仅使用边缘检测模块的输出作为预测，直接对测试图像输出提取边缘结果。

4. 实验结果

4.1. 实验设定

我们使用本文提出的光伏图像数据集进行实验，每张光伏图像拥有对应的边缘和语义分割标注。共包含 300 张光伏板图像，每张输入图的尺寸为 2560×1440 ，语义分割模块与 s2e 模块的损失函数权重分别为 $\lambda_1 = 0.5$ ， $\lambda_2 = 0.1$ 。为了对边缘检测模型的效果进行定量比较，本文首先将模型输出的边缘概率图转化为硬编码的边缘预测。考虑到边缘概率图中的元素代表图像上对应位置是边缘点的概率，取值范围为 0 到 1 之间，根据所设置的置信度阈值 η 将其转化为边缘预测图，该旅途上大于 η 的点被置为 1，小于 η 的点被置为 0，通过这种方式得到的二值边缘图用于后续与边缘标注之间进行计算。

本文首先计算检测结果的准确率 Precision 与召回率 Recall。其中，准确率表示正确预测为边缘的像素占全部预测为边缘的像素比例，召回率表示正确预测为边缘的像素占全部实际为边缘的像素比例。对准确率与召回率两个指标计算加权调和平均，F-Score：

$$F\text{-Score} = (1 + \beta^2) * \frac{\text{Precision} * \text{Recall}}{\beta^2 * \text{Precision} + \text{Recall}} \quad (9)$$

F-Score 值越高, 说明模型对于图像边缘的检测效果越好。在本文的实验中, 我们中立地综合考虑精确率和召回率, 即设置二者在评价指标中的权重相同, $\beta = 1$, 有:

$$F\text{-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

而为了排除阈值设置对于模型输出的影响, 本文进一步采用 ODS (Optimal Dataset Scale) 与 OIS (Optimal Image Scale) 指标对边缘检测结果进行评估, ODS 与 OIS 是两种不同的设置阈值 η 的方法。其中, ODS 为全局最佳方法, 是为一个数据集上的所有图像设置相同的阈值, 即选取能够使整个数据集上 F-Score 最大的阈值 η 应用于所有测试图片; 相比之下, OIS 为单图最佳方法, 这一方法在每张测试图像上针对性地选取使得该测试图片 F-Score 最大的阈值。

4.2. 实验结果与分析

根据从整体数据集中随即划分出来的一部分测试集上, 评估我们所提出的边缘检测模型的效果。可视化效果如图 5 所示, 图 5(a) 表示一张原始的测试图像, 图 5(b) 表示利用本文的基线模型-BDCN 方法得到的边缘图像, 图 5(c) 表示基于本文方法提取的边缘图像。可以看出, 本文方法得到的图像边缘能够与测试图像更加吻合, 得到的边缘线明显更加细致、精准; 非语义性的边缘, 如光伏板中间的白线也没有被提取出来。

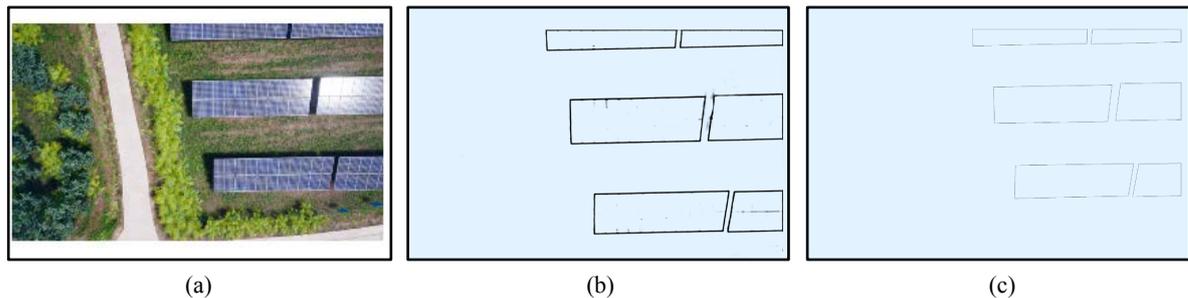


Figure 5. Result comparison of our method and BDCN. (a) Test image; (b) BDCN method result; (c) Our method result
图 5. 本文方法与 BDCN 方法结果比较。(a) 测试图像; (b) BDCN 方法结果; (c) 本文方法结果

此外, 我们利用量化指标来评估模型效果。我们将多人标注的边缘结果取均值作为标注, 并将模型得到的边缘与标注信息之间距离误差在 0.75 像素误差范围内的边缘均视为正确结果, 以进一步抵消人工标注所带来的误差。我们首先评估了 Sobel 和 Canny 这两种边缘检测算法生成的边缘, 但例如图 1 所示, Sobel 和 Canny 提取了大量错误的边缘结果, 生成的边缘图像较为混乱, 尤其是那些包含光照、树木、电线等场景的图像。Sobel 和 Canny 检测方法提取出的边缘 ODS 和 OIS 的精度均不足 5%, 说明传统的边缘检测算法很难在复杂场景下得到精确的结果。而下表 1 则列出了我们模型的消融实验结果。在作为基线的 BDCN 模型上, 我们依次添加多任务学习与多尺度特征融合。其中, 本文在多任务学习部分使用 Deeplab V3 作为语义分割任务的模型。

从实验结果可以看出, 本文提出的方法相较基线的 BDCN 方法, 显著提升了边缘检测结果。其中, 多任务学习方式在 ODS 与 OIS 两个指标上分别带来了 0.094 和 0.083 的精度提升, 充分说明语义分割任务对边缘检测的促进作用。同时, 多尺度特征融合模块的添加也能够提升模型性能, 相较于仅添加多任务学习的基线模型进一步提升了 0.035 和 0.034 的 ODS、OIS 精度。实验结果证明了在光伏电站这一复杂

场景中, 本文方法相比于基线 BDCN 模型的有效性, 且本文提出的多任务学习与多尺度特征融合都在一定程度上提升了边缘检测模型的能力。

Table 1. Comparison of ODS and OIS results

表 1. ODS 与 OIS 结果比较

	ODS	OIS
BDCN 基准	0.820	0.836
BDCN + 多任务学习	0.914	0.919
BDCN + 多任务学习 + 多尺度特征	0.949	0.953

5. 结束语

本文针对较为复杂的光伏电站场景, 提出了一种边缘检测算法。本文针对该场景中待检测物体可能较小、不相关噪声干扰等问题, 通过引入多尺度特征融合与多任务学习策略, 实现对光伏板边缘更加精细的检测效果。在本文提出的光伏板数据集上, 可视化实验结果与量化结果均证明了本文所提出方法的有效性。该方法可以进一步应用在光伏电站智能巡检等实际场景之中。

基金项目

基于无人机航拍图像的光伏电站智能巡检技术研究(B31205200009)。

参考文献

- [1] 陶洪久, 柳健, 田金文. 基于小波变换和数学形态学的遥感图像边缘检测[J]. 红外与激光工程, 2002, 31(2): 154-157.
- [2] 王鸿娟. 边缘检测在医学图像处理中的应用研究[J]. 自动化与仪器仪表, 2017(5): 140-141.
- [3] 赖思渝, 王宇峰, 王娟. 一种基于边缘检测的医学图像增强方法[J]. 电脑与信息技术, 2013(2): 17-20.
- [4] 霍桂利. 基于智能控制算法的自动驾驶系统优化研究[J]. 现代电子技术, 2019, 42(20): 177-180.
- [5] 俞涛. 基于深度学习的图像语义分割在自动驾驶中的应用[D]: [硕士学位论文]. 武汉: 华中科技大学, 2018.
- [6] Canny, J.F. (1987) A Computational Approach to Edge Detection. In: *Readings in Computer Vision*, Elsevier, Amsterdam, 184-203. <https://doi.org/10.1016/B978-0-08-051581-6.50024-6>
- [7] Taira, R.K., Iglesias, J.E. and Jahanshad, N. (2010) Characterizing Imaging Data. Springer US, Berlin. https://doi.org/10.1007/978-1-4419-0385-3_5
- [8] Xie, S. and Tu, Z. (2015) Holistically-Nested Edge Detection. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 1395-1403. <https://doi.org/10.1109/ICCV.2015.164>
- [9] He, J., Zhang, S., Yang, M., et al. (2020) Bi-Directional Cascade Network for Perceptual Edge Detection. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 3823-3832. <https://doi.org/10.1109/CVPR.2019.00395>
- [10] Hu, Z., Zhen, M., Bai, X., et al. (2020) JSENet: Joint Semantic Segmentation and Edge Detection Network for 3D Point Clouds. *European Conference on Computer Vision*, Heraklion, 5-11 September 2020, 222-239. https://doi.org/10.1007/978-3-030-58565-5_14
- [11] Liu, S., Johns, E. and Davison, A.J. (2018) End-to-End Multi-Task Learning with Attention. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 1871-1880. <https://doi.org/10.1109/CVPR.2019.00197>
- [12] Misra, I., Shrivastava, A., Gupta, A., et al. (2016) Cross-Stitch Networks for Multi-Task Learning. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 3994-4003. <https://doi.org/10.1109/CVPR.2016.433>
- [13] Liu, Y., et al. (2018) Richer Convolutional Features for Edge Detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **41**, 1939-1946.

- [14] Soria, X., Riba, E. and Sappa, A. (2020) Dense Extreme Inception Network: Towards a Robust CNN Model for Edge Detection. 2020 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, 1-5 March 2020, 1923-1932. <https://doi.org/10.1109/WACV45572.2020.9093290>
- [15] Yu, Z., Feng, C., Liu, M.Y., *et al.* (2017) CASENet: Deep Category-Aware Semantic Edge Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 5964-5973. <https://doi.org/10.1109/CVPR.2017.191>
- [16] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [17] Chen, L.C., Papandreou, G., Schroff, F., *et al.* (2017) Rethinking Atrous Convolution for Semantic Image Segmentation.