

# 基于多尺度特征编码和双重注意力融合的 绝缘子缺陷检测

李利荣<sup>1,2\*</sup>, 陈鹏<sup>1</sup>, 张云良<sup>1</sup>, 张开<sup>1</sup>, 熊炜<sup>1,2</sup>, 巩朋成<sup>1,2</sup>

<sup>1</sup>湖北工业大学电气与电子工程学院, 湖北 武汉 430064;

<sup>2</sup>新能源及电网装备安全监测湖北省工程研究中心, 湖北 武汉 430064

**摘要** 针对输电线路绝缘子缺陷检测准确率低和检测速度慢的问题,提出了一种基于多尺度特征编码和双重注意力融合的输电线路绝缘子缺陷检测方法。首先,为了使检测模型适应缺陷绝缘子特征尺度的多样性,编码网络采用 Res2Net50 提取更细粒度的特征,并在之后嵌入空洞空间金字塔池化模块实现多个尺度捕捉绝缘子及其缺陷的特征;其次,为了减少解码网络中特征信息的缺失,将主干网络的不同层特征与 efficient channel attention 注意力模块串联,并分别与经过 squeeze and excitation 注意力模块的各反卷积特征相加形成双重注意力融合。实验结果表明,所提方法的均值平均精度值约为 95.35%,每秒传输帧数约为 65.95,与其他方法相比,该方法对无人机绝缘子缺陷的准确检测具有一定的参考价值。

**关键词** 图像处理; 绝缘子缺陷检测; 空洞空间金字塔池化; 多尺度特征编码; 双重注意力融合

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2410008

## Insulator Defect Detection Based on Multi-Scale Feature Coding and Dual Attention Fusion

Li Lirong<sup>1,2\*</sup>, Chen Peng<sup>1</sup>, Zhang Yunliang<sup>1</sup>, Zhang Kai<sup>1</sup>, Xiong Wei<sup>1,2</sup>, Gong Pengcheng<sup>1,2</sup>

<sup>1</sup>School of Electrical and Electronic Engineering, Hubei University of Technology, Hubei 430064, Wuhan, China;

<sup>2</sup>Hubei Engineering Research Center of New Energy and Power Grid Equipment Safety Monitoring, Hubei 430064, Wuhan, China

**Abstract** Aiming at the problem of low detection accuracy and slow detection speed of insulator defects in transmission lines, a defect detection method for transmission line insulators based on multi-scale feature coding and double attention fusion is proposed. First, in order to adapt the detection model to the diversity of characteristic scales of defective insulators, the coding network uses Res2Net50 to extract more fine-grained features, and then embeds the atrous spatial pyramid pooling structure to capture the characteristics of insulators and their defects at multiple scales. Second, in order to reduce the lack of feature information in the decoding network, The different feature layers of the backbone network are connected in series with the efficient channel attention attention module, and they are added to the deconvolution features of the squeeze and excitation attention module to form a double attention fusion. Finally, Experiments show that the mean average precision index of the proposed method reaches about 95.35%, and the frames per second reaches about 65.95, and compared with other algorithms, this method has certain reference value for realizing the accurate detection of insulator defects of unmanned aerial vehicles.

**Key words** image processing; insulator defect detection; atrous spatial pyramid pooling; multi-scale feature coding; dual attention fusion

## 1 引言

绝缘子是高压输电系统用于电气隔离和固定导线

的重要设备,它在自然环境中长期受雨水、阳光、气候变化和化学腐蚀的影响,很容易受损,而绝缘子受损引起的故障会直接威胁到输电线路的稳定和安全<sup>[1-2]</sup>。

收稿日期: 2021-09-28; 修回日期: 2021-10-22; 录用日期: 2021-11-03

基金项目: 国家自然科学基金(62071172)、湖北省自然科学基金(2019CFB530)、新能源及电网装备安全监测湖北省工程研究中心开放研究基金(HBSKF202121)

通信作者: \*Rongli@hbut.edu.cn

因此需要对输电线路及绝缘子进行定期巡检。随着国民用电需求的不断增长,高压输电线路覆盖的范围也越来越广,传统的以人工跋山涉水及登塔巡检的方式已经满足不了当下长距离大范围的巡检任务。无人机结合计算机视觉算法在输电线路巡检任务上的逐渐应用,提高了输电线路的巡检效率、降低了输电线路巡检的成本和风险<sup>[3]</sup>。因此,利用无人机快速准确地对输电线路绝缘子进行检测对保障电力系统的运行具有十分重要的意义。

为了识别无人机采集图像中的缺陷绝缘子,传统的检测方法主要通过提取绝缘子的颜色、形状和纹理等常用特征来区分绝缘子和复杂背景<sup>[4]</sup>。例如:文献[5]提出了一种基于HSI模态强度的钢化玻璃绝缘子识别方法;文献[6]提出了一种基于轮廓特征和灰度相似度匹配的融合算法,实现了高精度的绝缘子轮廓提取和绝缘子片的精确分离,建立了基于绝缘子片间距和灰度相似度的绝缘子缺陷检测模型;文献[7]为了克服纹理不均匀带来的困难,提出了一种新的主动轮廓模型来从航拍图像中提取缺陷绝缘子。上述传统检测方法基于手工提取特征,效率低、泛化能力较差,而且只能在特定的检测条件下对绝缘子进行定位及故障检测<sup>[8]</sup>。

近十年来,随着计算机算力的不断提高,深度学习算法在目标检测领域取得了显著的成就。有些学者通过改进经典的目标检测算法来实现输电线路绝缘子缺陷检测<sup>[9-11]</sup>。例如:文献[12]基于单阶段SSD<sup>[13]</sup>检测算法设计了一种多级感知的航空图像绝缘子故障检测方法;文献[14]针对网络模型收敛速度较慢、目标回归框不够稳定、严重影响绝缘子定位精度和识别准确率的问题,提出了一种基于改进Faster RCNN<sup>[15]</sup>的绝缘子检测算法;文献[16]提出了一种改进的YOLOv3<sup>[17]</sup>模型,采用SPPNet<sup>[18]</sup>和多尺度预测网络提高绝缘子故障检测的准确性。总而言之,基于深度学习的绝缘子缺

陷检测方法与传统使用手工提取特征方法相比,前者可以从训练数据中自动学习,并在复杂环境中表现出更好的性能。在基于深度学习的方法中,两级检测器通常可以实现更高的检测精度,一级检测器对实时目标检测具有更强的适用性<sup>[19]</sup>。为了平衡检测精度和模型推理时间,基于Anchor-Free的目标检测算法摒弃了在特征图上生成大量锚框的思想,直接通过关键点回归目标框,极大地缩短了模型推理时间。比较典型的Anchor-Free算法有CornerNet<sup>[20]</sup>、ExtremeNet<sup>[21]</sup>、CenterNet<sup>[22]</sup>等。

为了改善检测模型对绝缘子缺陷检出率低、检测速度慢的问题,本文在具有编码解码结构的CenterNet目标检测网络基础上,提出了一种多尺度特征编码和双重注意力融合的输电线路绝缘子缺陷检测方法:首先,利用Res2Net50提取缺陷绝缘子的细粒度特征,并将其输出特征作为空洞空间金字塔池化(ASPP)模块的输入,得到多尺度特征,以改善网络对多尺度缺陷绝缘子的特征提取能力;然后,将特征提取网络Res2Net50的不同层特征分别串联efficient channel attention(ECA)模块形成跳跃支路,并与各反卷积层特征经过squeeze and excitation(SE)模块的输出特征相加,使预测网络的输入特征既具有高语义信息,又有底层的细节信息,从而获得更好的检测性能。

## 2 检测方法

### 2.1 CenterNet结构

CenterNet是一种基于ResNet18<sup>[23]</sup>、ResNet101、Hourglass104<sup>[24]</sup>和DLANet34<sup>[25]</sup>等4种特征提取网络设计的一种Anchor-Free的目标检测网络,如图1所示。CenterNet结构由编码网络、解码网络和检测网络等3部分构成。首先,编码网络对输入的图像进行特征提取,然后将具有高层语义信息的特征作为解码网络的输入,利用反卷积操作对特征图进行3次上采样

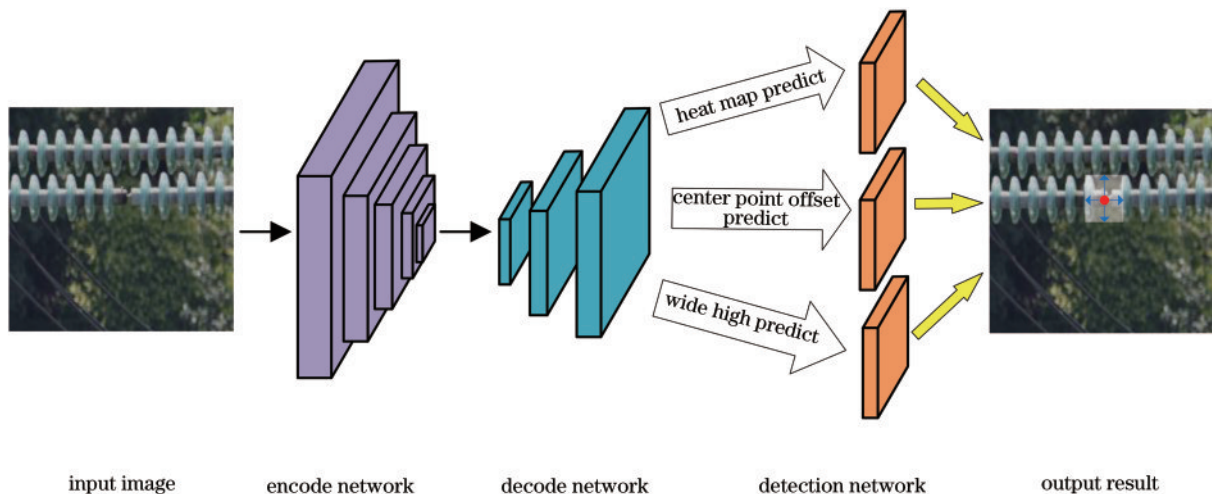


图1 CenterNet结构  
Fig. 1 CenterNet structure

得到高分辨率的特征图。检测网络由 3 个分支组成, 分别由一个  $3 \times 3$ 、 $1 \times 1$  的卷积组合共享同一个具有高分辨率的特征图, 3 个分支分别得到输出通道为目标类别数的热力图、目标中心距离热力峰值的偏移、物体宽高信息。最后提取热力图峰值点、预测的中心点偏移及物体宽高等信息, 输出检测结果。

### 2.2 所提算法结构

所提算法是在 CenterNet 的基础进行改进的, 如图 2 所示。编码网络使用 Res2Net50<sup>[26]</sup> 作为主干网络对输入的图像进行细粒度的特征提取, 并针对输电线路绝

缘子及其缺陷尺度变化大的问题, 在主干网络提取的高语义特征层后增设 ASPP<sup>[27]</sup> 模块, 通过选择合适的空洞率提取多个尺度的空间信息以增大其感受野。解码网络首先在每一个反卷积操作之后嵌入 SE<sup>[28]</sup> 模块对特征图的权值进行重新标定, 抑制无用的特征信息。同时将编码网络各特征层的输出  $L_1$ 、 $L_2$ 、 $L_3$  经过高效注意力 ECA<sup>[29]</sup> 模块提高通道特征的注意力, 依次得到  $Y_1$ 、 $Y_2$ 、 $Y_3$ , 然后将反卷积特征  $D_1$ 、 $D_2$ 、 $D_3$  通过 SE 模块重新标定后与特征  $Y_1$ 、 $Y_2$ 、 $Y_3$  相加形成双重注意力融合的解码结构, 最后通过预测网络的 3 个分支完成目标检测任务。

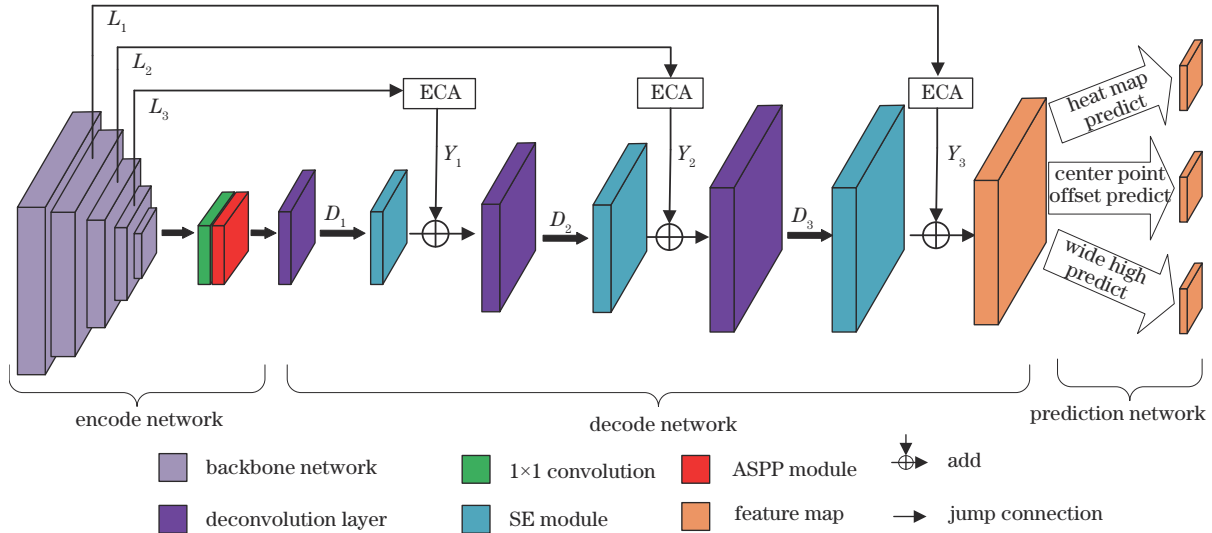


图 2 所提算法结构  
Fig. 2 Structure of proposed algorithm

### 2.3 多尺度特征编码

传统的特征提取网络主要通过堆叠更多的卷积层或者使用多个小卷积核组合的方式来扩大感受野范围。直接堆叠大量的卷积层会导致每个特征层的感受野相对固定, 小卷积核组合虽然可以大幅降低模型的参数量和复杂度, 但是不能完全凸显特征信息。为了使检测模型适应绝缘子及其缺陷的特征尺度, 本研究采用 Res2Net50 与 ASPP 串联的方法组成多尺度特征编码结构。

#### 2.3.1 Res2Net50

基于深度学习的目标检测算法通常使用 ResNet 系列作为主干特征提取网络, ResNet 因在 Bottleneck 模块 [如图 3(a) 所示] 中采用短路连接的方式获得更深层次的网络结构并缓解梯度消失而得到广泛应用。考虑到 ResNet18 的网络层数较少, 提取高语义信息的能力不足, 虽然 ResNet101 通过短路连接允许不同的卷积算子组合, 但是会产生大量的等价特征。本研究使用 Res2Net50 作为主干特征提取网络, Res2Net50 将 ResNet 残差块中的  $3 \times 3$  卷积都替换为如图 3(b) 中虚线框所示的结构。具体地说: 首先将输入的特征图经过  $1 \times 1$  卷积后按通道数均匀地分成 4 组特征子图;

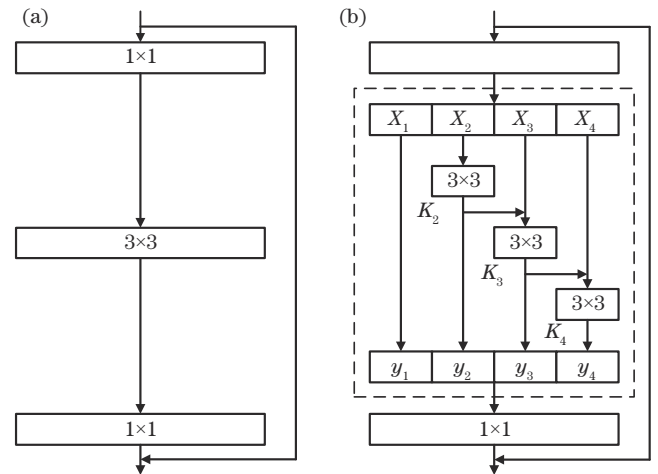


图 3 Bottleneck block 与 Res2Net module 结构图。  
(a) Bottleneck block; (b) Res2Net module

Fig. 3 Schematic diagrams of Bottleneck block and Res2Net Module. (a) Bottleneck block; (b) Res2Net module

除了  $X_1$  组特征子图没有进行  $3 \times 3$  卷积操作之外, 其他组特征子图都依次进行  $3 \times 3$  卷积之后与下一组特征子图进行拼接操作并再次进行  $3 \times 3$  卷积, 直到这 4 组特征子图都被依次处理完; 最后将所有具有不同感受

野的特征子图组拼接起来,通过 $1 \times 1$ 的卷积实现特征融合。综上所述,Res2Net50以分组分层的类残差方式连接不同的卷积核,使每个残差块在不明显增加参数的同时确保提取到更充足的语义特征。

### 2.3.2 嵌入ASPP模块

主干网络提取特征后,通过 $1 \times 1$ 卷积将具有高语义信息的特征降维再作为ASPP模块(如图4所示)的输入实现多个尺度捕捉绝缘子及其缺陷的特征信息,

ASPP模块由全局平均池化、 $1 \times 1$ 卷积和3个不同空洞率的 $3 \times 3$ 卷积并行输出,随后将5个卷积的输出拼接并通过 $1 \times 1$ 卷积融合输出具有多尺度信息的特征图。本研究选择空洞率分别为6、12、18的卷积组合与主干网络构成多尺度特征编码网络,这样的组合既可以平衡网络对于输电线路绝缘子及其缺陷的敏感程度,又可以提高对输电线路不同尺度绝缘子及其缺陷的检测性能。

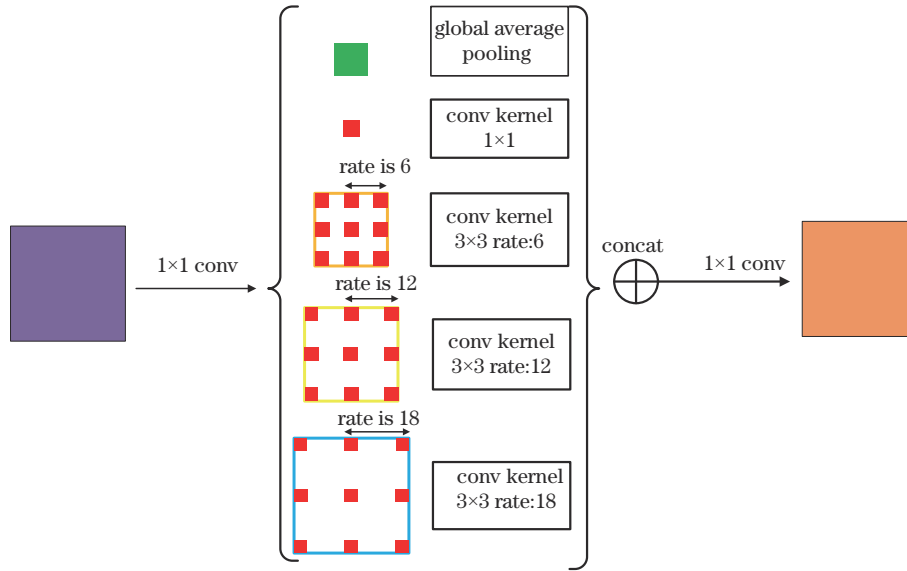


图4 ASPP模块

Fig. 4 ASPP module

### 2.4 双重注意力融合

为了减少解码网络中特征信息的损失,将主干网络的 $L_1$ 、 $L_2$ 、 $L_3$ 特征串联输入高效的通道注意力机制ECA模块,然后与 $D_1$ 、 $D_2$ 、 $D_3$ 反卷积特征经过SE模块后输出的特征相加形成双重注意力融合,如图2所示。

具体融合操作如图5所示:第1条支路SE模块支路,首先将解码网络中反卷积输出的特征 $D_i(i=1,2,3)$ 通过卷积变换 $F_{tr}(\cdot, \theta)$ 得到具有全局特征信息的特征图 $U$ ,然后将特征图 $U$ 通过聚合函数 $F_{sq}(\cdot)$ 得到一个具有局部聚合的 $1 \times 1 \times C_2$ 特征向量,最后使用 $F_{ex}(\cdot, u)$ ,

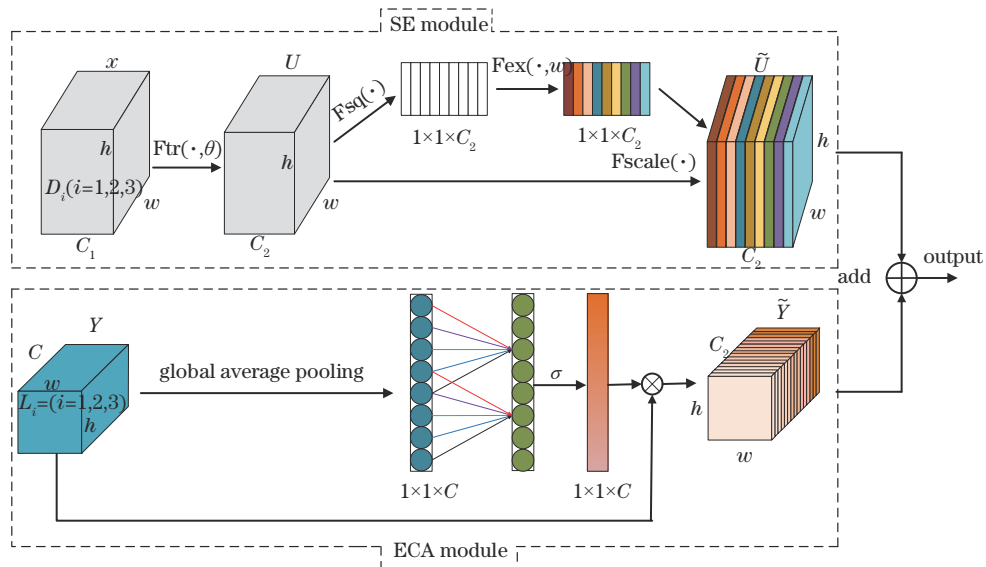


图5 双重注意力融合示意图

Fig. 5 Schematic diagram of dual attention fusion

w)对局部聚合特征经过全连接、激活操作得到具有通道权重的 $1 \times 1 \times C_2$ 特征向量,并与特征图 $U$ 使用矩阵相乘输出 $w \times h \times C_2$ 的特征图;第2条支路将来自主干网络的 $L_i(i=1,2,3)$ 层特征作为ECA模块的输入,首先通过全局平均池化得到 $1 \times 1 \times C$ 的特征向量,然后通过自适应地选择一维卷积核大小的方法获取特征的局部信息,最后通过Sigmoid激活函数得到 $1 \times 1 \times C$ 的特征通道权重向量,并与ECA的输入特征按对应通道相乘输出 $w \times h \times C_2$ 的特征图。将两条支路的输出特征融合得到具有低维细节和高维语义的特征图,使模型更加凸显有用的通道特征,提高对输电线路绝缘子及其缺陷的检测精度。

### 2.5 损失函数

所提模型训练的损失函数由热力图关键点损失、中心点偏移损失、尺寸大小损失等3部分构成。首先,输入图像 $I \in \mathbf{R}^{W \times H \times 3}$ ( $W$ 为图像的宽度、 $H$ 为图像的高度)的目的是得到特征图 $\hat{Y}_{x,y,c} \in [0, 1]^{\frac{W}{T} \times \frac{H}{T} \times C}$ ( $T$ 为下采样倍数、 $C$ 为关键点类别数),然后通过预测网络的3个分支回归物体的类别、中心点偏移和尺寸信息。 $\hat{Y}_{x,y,c} = 1$ 代表存在或检测到目标,相反 $\hat{Y}_{x,y,c} = 0$ 代表检测到背景。对于每一个目标类别 $C$ 的关键点 $p \in \mathbf{R}^2$ , $T$ 倍下采样数可表示为 $\tilde{p} = [\frac{p}{T}]$ 。然后,利用高斯核 $Y_{xyc}$ 将所有的真实关键点投射到关键点热图上。

$$Y_{xyc} = \exp\left[-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\sigma_p^2}\right], \quad (1)$$

式中: $\sigma_p$ 为尺寸自适应标准差。

为了监督真实关键点生成热力图,使用Focal损失作为关键点损失函数:

$$L_k = \frac{-1}{N} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \lg \hat{Y}_{xyc}, & \hat{Y}_{xyc} = 1 \\ (1 - \hat{Y}_{xyc})^\beta (\hat{Y}_{xyc})^\alpha \lg(1 - \hat{Y}_{xyc}), & \hat{Y}_{xyc} \neq 1 \end{cases}, \quad (2)$$

式中: $N$ 为图像中的类别数;超参数 $\alpha$ 和 $\beta$ 分别取2和4。

为了使生成的关键点中心与真实目标中心点匹配,使用通道数为2的特征图预测一个局部偏移 $\hat{O} \in \mathbf{R}^{\frac{W}{T} \times \frac{H}{T} \times 2}$ ,目标中心点的偏移损失函数为

$$L_{\text{off}} = \frac{1}{N} \sum_p |\hat{O} - (\frac{p}{T} - \tilde{p})|, \quad (3)$$

式中: $\hat{O}$ 为模型预测的中心点偏移; $\frac{p}{T} - \tilde{p}$ 为实际中心点偏移; $p$ 为实际目标中心; $\tilde{p}$ 为预测的中心点坐标。

取目标 $k$ 的边界框对侧角点坐标 $(x_1^{(k)}, y_1^{(k)}, x_2^{(k)}, y_2^{(k)})$ ,将边界框的坐标转换为目标的尺寸 $S_k = (x_1^{(k)} - x_2^{(k)}, y_1^{(k)} - y_2^{(k)})$ ,为了减少计算量,同样使用通道数为2的特征图对所有目标种类预测出尺寸 $\hat{S}_{pk} \in \mathbf{R}^{\frac{W}{T} \times \frac{H}{T} \times 2}$ ,使用 $L_{\text{size}}$ 函数监督模型训练过程中的尺

寸大小损失。

$$L_{\text{size}} = \frac{1}{N} \sum_{k=1}^N |\hat{S}_{pk} - S_k|, \quad (4)$$

式中: $\hat{S}_{pk}$ 为预测的物体尺寸; $S_k$ 为图像的真实尺寸。

则整个网络的损失函数为

$$L_{\text{det}} = L_k + \lambda_{\text{size}} L_{\text{size}} + \lambda_{\text{off}} L_{\text{off}}, \quad (5)$$

式中:权重系数 $\lambda_{\text{size}}$ 和 $\lambda_{\text{off}}$ 分别取0.1和1。

## 3 实验过程及对比实验分析

### 3.1 实验过程

本研究所有实验基于的软件环境是Python 3.6、Pycharm社区版和Pytorch 1.2深度学习框架,均在一台配备Intel i7-8700 3.2 GHz CPU和一个16 GB运行内存的NVIDIA RTX 2080Ti GPU的工作站进行。为了使时间和资源都能得到充分利用,采用冻结与解冻的方式进行训练:首先将主干网络冻结,使用迁移学习的方式对网络训练50个Epoch;然后解冻主干网络,与解码网络和预测网络一起端到端训练150个Epoch得到检测模型,训练损失和验证曲线如图6所示。由于输入的图片尺寸太大,消耗大量的显存,将批处理大小(batch size)设为4。选用Adam网络优化器,网络初始学习率设为 $1 \times 10^{-3}$ 。

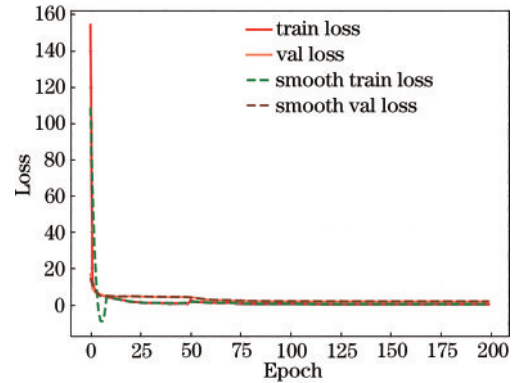


图6 损失曲线

Fig. 6 Loss curve

### 3.2 数据集

实验采用中国电力线绝缘子数据集,它包含无人机拍摄的600张正常绝缘子图像及248张有缺陷的绝缘子图像。为了解决网络训练数据包含有缺陷绝缘子的图像数量有限且训练数据规模不均衡的问题,采用数据增强的方法将数据集扩大到1509张图像,其中有缺陷的绝缘子图像扩大到754张,所有图像像素大小均为 $512 \times 512$ ,数据集示例如图7所示。为了使网络成功训练,利用LabelImg工具标注每幅图像绝缘子及其缺陷绝缘子的真实框标签,并将所有标签文件与图像整理成PASCAL VOC数据集格式,选择其中1207幅图像为训练集(其中200幅为验证集),302幅图像为测试集。真实框标签示例如图8所示。



图 7 数据集示例  
Fig. 7 Sample dataset



图 8 真实框标签示例  
Fig. 8 Example of real box labels

### 3.3 评价指标

采用精确度( $P$ )、召回率( $R$ )、平均精度( $AP$ )、均值平均精度( $mAP$ )、F1 分数( $F1$ -score)和每秒传输帧数( $FPS$ )等 6 个指标评价模型性能。二元分类问题中真阳性( $TP$ )、假阳性( $FP$ )、真阴性( $TN$ )、假阴性( $FN$ )的定义如表 1 所示,其中“1”和“0”分别代表正样本和负样本。

表 1 TP、TN、FP、FN 的定义

Table 1 Definitions of TP, TN, FP, and FN

Real category	Predictive value	Definition
1	1	TP
1	0	FP
0	1	TN
0	0	FN

$P$ 为模型预测为正样本且真实类别也为正样本的数量占模型预测为正样本数量的比例。 $R$ 为模型预测为正样本且真实类别也是正样本的数量占真实类别是正样本数量的比例。以 $P$ 作 $y$ 轴, $R$ 作 $x$ 轴,得到 $P-R$ 曲线, $P-R$ 曲线围成的面积即为AP。 $mAP$ 为所有类别AP的平均值。 $F1$ 分数为精度与召回率的调和平均值。 $FPS$ 为设备每秒传输帧数,对于嵌入式设备来说通常FPS达到20就可以满足实时的要求。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (6)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (7)$$

$$P_{AP} = \int_0^1 P(R) dR, \quad (8)$$

$$P_{mAP} = \frac{P_{AP1} + P_{AP2} + \dots + P_{APn}}{n}, \quad (9)$$

$$S_{F1} = 2 \times \frac{P \times R}{P + R}, \quad (10)$$

式中: $n$ 为目标类别数。

### 3.4 不同主干网络的比较分析

表2列出了基于CenterNet目标检测框架使用5种不同主干网络生成模型的评价结果。从表2的实验数据可以看出:以Res2Net50为主干网络的 $mAP$ 指标达到88.27%,与ResNet18、ResNet50、ResNet101、DLANet34为主干网络相比,分别高出41.9个百分点、2.68个百分点、11.1个百分点、10.16个百分点;特别地,以Res2Net50为主干网络的模型对缺陷绝缘子的AP50、 $P$ 和 $R$ 指标相较于其他4种模型是最高的,这表明Res2Net50以分组分层的类残差方式连接不同的卷积核相较于其他4种主干网络对缺陷绝缘子具有更强的特征提取能力。因此,本研究选择Res2Net50作为特征提取网络。

表2 不同主干网络实验结果

Table 2 Experimental results of different backbone networks

unit: %

Algorithm framework	Backbone network	AP50		$P$		$R$		mAP
		Normal	Defect	Normal	Defect	Normal	Defect	
CenterNet	ResNet18	80.80	14.95	96.86	50.00	41.48	0.79	47.87
	ResNet50	<b>94.72</b>	76.46	94.20	89.47	84.08	53.54	85.59
	ResNet101	91.65	62.70	98.86	84.00	77.80	33.07	77.17
	DLANet34	93.61	62.61	<b>99.41</b>	81.40	75.56	27.56	78.11
	Res2Net50	94.26	<b>82.29</b>	95.93	<b>91.09</b>	<b>84.53</b>	<b>72.44</b>	<b>88.27</b>

### 3.5 消融实验分析

以CenterNet为基础框架,依次使用Res2Net50、Res2Net50+ASPP、Res2Net50+ASPP+ECA、Res2Net50+ASPP+ECA+SE的组合进行实验。消融实验的数据如表3、表4所示。表3、表4中第1组实验是在使用Res2Net50作为主干网络的基础上只添加侧向连接时生成模型的指标,与表2中选择Res2Net50无侧向连接的模型相比,添加跳跃支路将 $mAP$ 提高了1.5个百分点,即添加跳跃支路融合了绝缘子和缺陷绝缘子特征减少了特征信息的丢失,提高了绝缘子及缺陷绝缘子预测的准确性。第2组实验数据表明,在第1组实验的基础上添加ASPP模块能有效地多尺度捕捉绝缘子

及其缺陷的特征信息, $mAP$ 相比于第1组提升了2.19个百分点。第3组实验在第2组实验的基础上,在侧向支路中添加高效的通道注意力机制ECA模块,然后与解码网络的反卷积特征层进行特征融合,增强了反卷积后特征的通道信息,使得缺陷绝缘子的AP指标相较于第2组实验提升了3.69个百分点。第4组实验是所提算法,其AP50、 $P$ 、 $R$ 、 $F1$ 和 $mAP$ 评价指标相较于其他组实验都是最高的。

另外,在基于Res2Net50和ASPP构成多尺度特征编码结构的情况下还针对两条支路添加ECA和SE的情况进行了对比实验,实验数据如表5所示。第1支路为编码网络的各层特征 $L_i$ 支路,第2支路为解码网

表3 消融实验

Table 3 Ablation experiment

unit: %

No.	Res2Net50	ASPP	ECA	SE	$P$		$R$	
					Normal	Defect	Normal	Defect
1	✓				98.68	92.22	83.63	65.35
2	✓	✓			97.16	92.52	84.53	77.95
3	✓	✓	✓		95.73	94.50	85.43	81.10
4	✓	✓	✓	✓	<b>97.26</b>	<b>96.36</b>	<b>87.44</b>	<b>83.46</b>

表 4 消融实验  
Table 4 Ablation experiment

No.	Res2Net50	ASPP	ECA	SE	AP50		F1		mAP / %
					Normal / %	Defect / %	Normal	Defect	
1	✓				93.29	86.24	0.91	0.76	89.77
2	✓	✓			95.33	88.58	0.90	0.85	91.96
3	✓	✓	✓		94.73	92.27	0.90	0.87	93.50
4	✓	✓	✓	✓	<b>95.88</b>	<b>94.81</b>	<b>0.92</b>	<b>0.89</b>	<b>95.35</b>

表 5 两条支路添加不同注意力对比实验  
Table 5 Comparison experiment of adding different attention on two branches unit: %

The first branch	The second branch	AP50		<i>P</i>		<i>R</i>		mAP
		Normal	Defect	Normal	Defect	Normal	Defect	
ECA	ECA	95.28	89.70	97.74	93.33	87.44	77.17	92.49
ECA	SE	95.88	94.81	97.26	96.36	87.44	83.46	95.35

络反卷积后的特征  $D_i$  支路,当这两条支路同时添加高效的注意力 ECA 模块时的 mAP 指标要比第 1 支路串联 ECA 模块、第 2 支路串联 SE 模块高出 2.86 个百分点,这表明将 ECA 模块和 SE 模块融合形成双重注意力对模型检测性能的提升更明显。

### 3.6 不同算法比较

为了更好地验证所提算法的检测性能,将所提算法与 5 种基于深度学习的目标检测框架 SSD、RetinaNet、Faster RCNN、YOLOv3、CenterNet 在相同的数据集下进行了对比实验,结果如表 6 所示。从表 6 可以看出:所提的 mAP 指标高于单阶段 SSD、RetinaNet 和 YOLOv3 检测算法,分别高出 5.4 个百分

点、7.81 个百分点和 4.67 个百分点;与两阶段检测算法 Faster RCNN 相比,虽然 Faster RCNN 检测模型把所有的正类都找出来的能力很强(召回率越高,将所有正类都检测出的能力越强),但是 mAP 指标较低,所提算法的 FPS 是 Faster RCNN 的 2 倍多;将所提算法与 CenterNet 相比,除 FPS 外的 5 个评价指标均高于 CenterNet,特别地,缺陷绝缘子的 AP50、*P*、*R* 指标分别高出了 18.35 个百分点、6.89 个百分点、29.92 个百分点,mAP 高出了 9.76 个百分点。综上所述,所提算法的检测精度和检测速度相较于单、两阶段检测算法具有一定的优势。

表 6 不同算法的对比  
Table 6 Comparison of different algorithms unit: %

Algorithm framework (backbone)	AP50		<i>P</i>		<i>R</i>		mAP	FPS
	Normal	Defect	Normal	Defect	Normal	Defect		
SSD(VGG16)	91.02	88.87	93.75	92.86	87.5	88.67	89.95	58.62
RetinaNet(ResNet50)	88.70	86.38	95.91	88.03	78.92	84.1	87.54	44.30
FasterRCNN(VGG16)	<b>96.25</b>	70.49	73.47	42.91	<b>96.86</b>	<b>95.28</b>	83.37	29.16
YOLOv3(Darknet53)	93.90	87.45	91.61	91.27	93.05	90.55	90.68	75.63
CenterNet(ResNet50)	94.72	76.46	94.20	89.47	84.08	53.54	85.59	<b>97.51</b>
Proposed algorithm	95.88	<b>94.81</b>	<b>97.26</b>	<b>96.36</b>	87.44	83.46	<b>95.35</b>	65.95

### 3.7 检测结果可视化

将所提基于多尺度特征编码和双重注意力融合的检测模型在测试集上进行测试,检测结果可视化如图 9 所示。从图 9(a)可以看出,检测模型对于在复杂

背景和由视距差而成像的不同尺寸绝缘子检测准确性能达到较好的效果。从图 9(b)可以看出,检测模型对被杆塔遮挡的绝缘子和缺陷绝缘子检测准确性也能达到不错的效果。



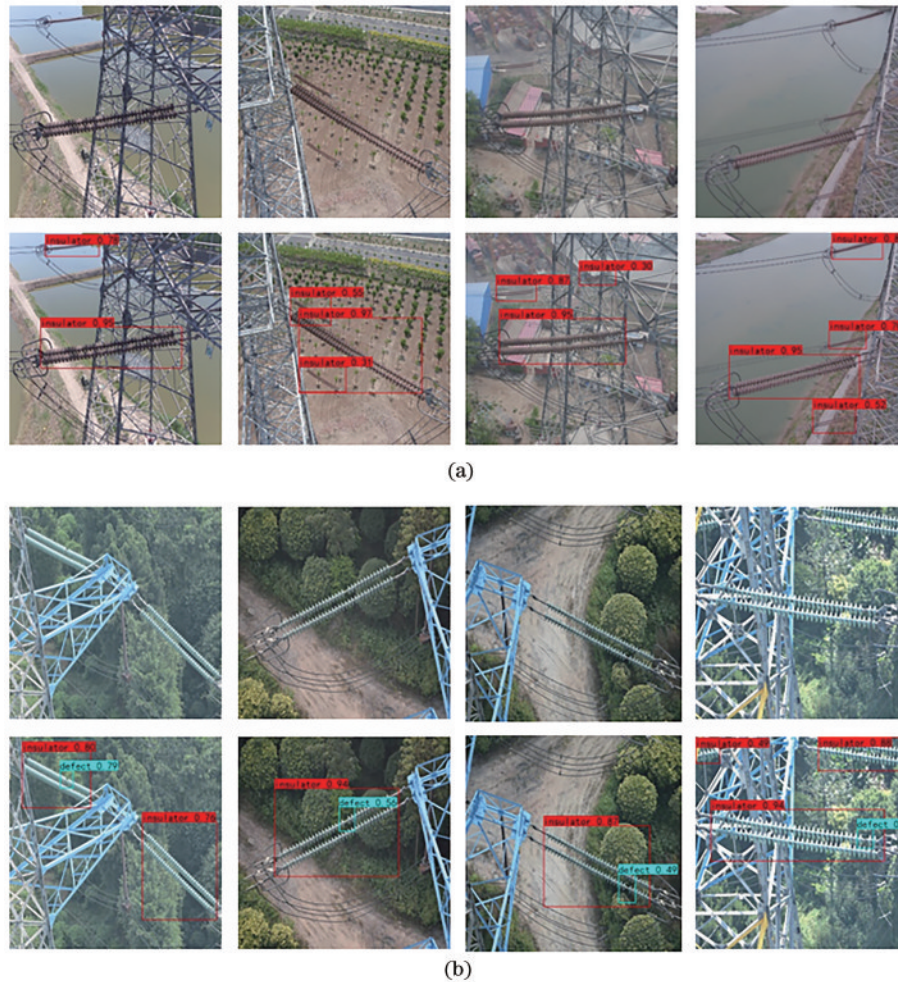


图9 检测结果可视化。(a)绝缘子检测结果可视化;(b)绝缘子及缺陷绝缘子的检测结果可视化

Fig. 9 Visualization of detection results. (a) Visualization of insulator detection results; (b) visualization of test results of insulators and defective insulators

## 4 结 论

提出了一种基于多尺度特征编码和双重注意力融合的输电线路绝缘子缺陷检测模型。输入输电线路巡检图像或视频,通过基于多尺度特征编码网络、双重注意力融合的解码网络和预测网络实现绝缘子缺陷检测。使用 Res2Net50 作为主干特征提取网络,并在特征提取网络之后嵌入 ASPP 模块构成多尺度特征编码网络。将特征提取网络的不同层特征分别串联 ECA 模块形成的跳跃支路与各反卷积层特征经过 SE 模块的输出相加构成双重注意力融合解码结构。最后,进行了对比和消融实验。实验结果表明,所提算法的 mAP 评价指标在对比和消融实验中是最高的,达到 95.35%,相较于 SSD、RetinaNet、Faster-RCNN、YOLOv3 和 CenterNet 分别高出 5.4 个百分点、7.81 个百分点、11.98 个百分点、4.47 个百分点和 9.76 个百分点, FPS 达到 65.95,满足实时检测要求。所设计的检测模型在输电线路绝缘子缺陷检测中具有一定的应用价值。

## 参 考 文 献

- [1] Park K C, Motai Y, Yoon J R. Acoustic fault detection technique for high-power insulators[J]. IEEE Transactions on Industrial Electronics, 2017, 64(12): 9699-9708.
- [2] Lei X S, Sui Z H. Intelligent fault detection of high voltage line based on the Faster R-CNN[J]. Measurement, 2019, 138: 379-385.
- [3] Liang H G, Zuo C, Wei W M. Detection and evaluation method of transmission line defects based on deep learning[J]. IEEE Access, 2020, 8: 38448-38458.
- [4] Miao X R, Liu X Y, Chen J, et al. Insulator detection in aerial images for transmission line inspection using single shot multibox detector[J]. IEEE Access, 2019, 7: 9945-9956.
- [5] Zhang X Y, An J B, Chen F M. A simple method of tempered glass insulator recognition from airborne image [C]//2010 International Conference on Optoelectronics and Image Processing, November 11-12, 2010, Haikou, China. New York: IEEE Press, 2010: 127-130.
- [6] Tan P, Li X F, Xu J M, et al. Catenary insulator defect detection based on contour features and gray similarity

- matching[J]. Journal of Zhejiang University-SCIENCE A, 2020, 21(1): 64-73.
- [7] Wu Q G, An J B. An active contour model based on texture distribution for extracting inhomogeneous insulators from aerial images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(6): 3613-3626.
- [8] Tao X, Zhang D P, Wang Z H, et al. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(4): 1486-1498.
- [9] 王勃凡, 赵海涛. 基于径向基激活函数的高光谱小目标检测[J]. 光学学报, 2021, 41(23): 2311001.  
Wang B F, Zhao H T. Small object detection in hyperspectral images based on radial basis activation function[J]. Acta Optica Sinica, 2021, 41(23): 2311001.
- [10] 梁霄, 李家伟, 赵小龙, 等. 基于深度学习的红外目标成像液位检测方法[J]. 光学学报, 2021, 41(21): 2110001.  
Liang X, Li J W, Zhao X L, et al. Infrared target imaging liquid level detection method based on deep learning[J]. Acta Optica Sinica, 2021, 41(21): 2110001.
- [11] 胡杰, 刘汉, 徐文才, 等. 基于三维激光雷达的道路障碍物目标位姿检测算法[J]. 中国激光, 2021, 48(24): 2410001.  
Hu J, Liu H, Xu W C, et al. Position detection algorithm of road obstacles based on 3D LiDAR[J]. Chinese Journal of Lasers, 2021, 48(24): 2410001.
- [12] Jiang H, Qiu X J, Chen J, et al. Insulator fault detection in aerial images based on ensemble learning with multi-level perception[J]. IEEE Access, 2019, 7: 61797-61810.
- [13] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [14] 马耀名, 张雨. 基于改进更快速区域卷积神经网络的绝缘子检测算法[J/OL]. 计算机应用: 1-10 [2021-10-21]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210531.0944.002.html>.  
Ma Y M. Insulator detection algorithm based on improved and faster region convolutional neural network [J/OL]. Journal of Computer Applications: 1-10 [2021-10-21]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210531.0944.002.html>.
- [15] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [16] Liu J J, Liu C Y, Wu Y Q, et al. An improved method based on deep learning for insulator fault detection in diverse aerial images[J]. Energies, 2021, 14(14): 4365.
- [17] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2021-09-16]. <https://arxiv.org/abs/1804.02767>.
- [18] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2015, 37(9): 1904-1916.
- [19] Wu X W, Sahoo D, Hoi S C H. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020, 396: 39-64.
- [20] Law H, Deng J. CornerNet: detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3): 642-656.
- [21] Zhou X Y, Zhuo J C, Krähenbühl P. Bottom-up object detection by grouping extreme and center points[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 850-859.
- [22] Zhou X Y, Wang D Q, Krähenbühl P. Objects as points [EB/OL]. (2019-04-25)[2021-09-16]. <https://arxiv.org/abs/1904.07850>.
- [23] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [24] Bao W X, Yang Y P, Liang D, et al. Multi-residual module stacked hourglass networks for human pose estimation[J]. Journal of Beijing Institute of Technology (English Edition), 2020, 29(1): 110-119.
- [25] Feng Z Y, Jin L W, Tao D P, et al. DLANet: a manifold-learning-based discriminative feature learning network for scene classification[J]. Neurocomputing, 2015, 157: 11-21.
- [26] Gao S H, Cheng M M, Zhao K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.
- [27] 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.
- [28] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [29] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.