

# 基于注意力机制的变电站作业场景三维目标检测

高伟<sup>1</sup>, 何搏洋<sup>1</sup>, 张婷<sup>2</sup>, 郭美青<sup>2</sup>, 刘军<sup>2</sup>, 王慧民<sup>2</sup>, 张兴忠<sup>2\*</sup>

<sup>1</sup>国网山西省电力公司互联网部, 山西 太原 030021;

<sup>2</sup>太原理工大学软件学院, 山西 晋中 030600

**摘要** 在变电站场景中, 作业人员与危险设备的空间距离感知是安全管控任务的基本问题。随着激光雷达和三维(3D)视觉理论的发展, 3D点云目标检测可为下游空间距离度量任务提供必要的技术支撑。针对变电站场景下背景复杂、设备遮挡等因素引起的目标检测不准的问题, 基于PointNet++模型, 在局部特征提取阶段引入改进的注意力模块, 提出了一种适用于变电站作业场景的3D目标检测网络PowerNet。首先经过两级局部特征提取, 获取每个局部区域中的细粒度特征; 其次通过mini-pointnet将所有局部特征编码成特征向量, 得到全局特征; 最后由全连接层输出预测结果。考虑到变电站点云数据中前景点与背景点数量差距较大, PowerNet采用Focal损失计算分类损失, 使网络更加关注前景点特征信息。在自建数据集上的实验结果表明, PowerNet的均值平均精度(mAP)值达到0.735, 高于其他模型, 可直接在下游安全管控任务中应用。

**关键词** 图像处理; 变电站作业场景; 三维目标检测; 点云; 通道方向注意力; 点方向注意力

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP202259.2210010

## Three-Dimensional Object Detection in Substation Operation Scene Based on Attention Mechanism

Gao Wei<sup>1</sup>, He Boyang<sup>1</sup>, Zhang Ting<sup>2</sup>, Guo Meiqing<sup>2</sup>, Liu Jun<sup>2</sup>, Wang Huimin<sup>2</sup>, Zhang Xingzhong<sup>2\*</sup>

<sup>1</sup>Internet Department, State Grid Shanxi Electric Power Company, Taiyuan 030021, Shanxi, China;

<sup>2</sup>College of Software, Taiyuan University of Technology, Jinzhong 030600, Shanxi, China

**Abstract** The perception of the spatial distance between operators and dangerous equipment is a basic safety management and control task issue in a substation scene. With the advancement of lidar and three-dimensional (3D) vision theory, 3D point cloud target detection can provide necessary technical assistance for downstream spatial distance measurement tasks. Aiming at the problem of inaccurate target detection caused by factors such as complex background and equipment occlusion in the substation scene, based on the PointNet++ model, an improved attention module is introduced in the local feature extraction stage, and a 3D object detection network PointNet suitable for substation operation scene is proposed. First, the network undergoes a two-level local feature extraction to obtain fine-grained features in each local area, then encodes all local features into feature vectors using a mini-pointnet to obtain global features, and finally passes through the fully connected layer to predict the results. Considering the large gap between the number of front and background points in the cloud data of substation sites, this study calculates the classification loss using focal loss to make the network pay more attention to the feature information of the front points. Experiments on the self-built dataset show that the PowerNet has a mean average precision (mAP) value of 0.735, which is greater than previous models and can be directly applied to downstream security management and control tasks.

**Key words** image processing; substation operation scene; three-dimensional object detection; point cloud; channel-direction attention; point-direction attention

收稿日期: 2021-08-23; 修回日期: 2021-10-01; 录用日期: 2021-10-27

基金项目: 山西省重点研发计划项目(201803D31041)、国网山西省电力公司信息化项目(B1051C200016)

通信作者: \*zhangxzsubmit@126.com

# 1 引言

变电站是连接电力传输的重要环节<sup>[1]</sup>,其运行稳定性直接影响到电网调度的可靠性和安全性<sup>[2]</sup>。对变电站场景内电力设备进行巡检是保证变电站正常工作的必要手段。统计分析表明,变电站场景内存在变压器等多种危险设备,造成变电站内发生人身安全事故的原因大多是工作人员误入带电间隔、走错工作区域或接触带电设备<sup>[3]</sup>。因此,在变电站巡检任务中采取有效的安全距离管控方法就显得尤为关键。

现有的变电站安全监控主要有视频监控和变电站虚拟三维(3D)可视化安全管控两种模式<sup>[3]</sup>。视频监控仅依赖二维(2D)视频流,无法度量空间中三维目标之间的距离。文献[3-5]将虚拟三维可视化技术应用于变电站作业场景,但此类方法通常需要使用差分定位或其他定位技术对作业人员的位置信息进行校正和解算,计算过程复杂且定位效果受多种因素影响。考虑到变电站作业场景内的安全距离管控依赖于三维目标的准确定位,而激光雷达扫描得到的三维点云数据包含了丰富准确的三维位置信息<sup>[6]</sup>,所以本实验组主要研究如何利用激光点云数据检测变电站作业场景中的三维目标,为下游变电站安全管控任务提供技术支撑。

基于激光点云数据的三维目标检测算法包括基于视图的方法、基于体素的方法及基于点的方法。基于视图的方法采用不同的点云投影方法将三维点云转换为二维视图,采用二维目标检测网络<sup>[7-8]</sup>实现目标检测,再使用位置维度回归恢复目标在三维空间中的几何姿态,具体可参考文献[9-12]。该类方法在降低复杂度的同时也造成了较大程度的信息损失,不利于变电站作业场景中的精确检测。基于体素的方法直接在三维空间中点云划分为规则的立方体,将其体素化来编码点云的三维特征,并将二维卷积神经网络拓展到三维卷积神经网络,具体可参考文献[13-16]。但随着空间复杂度的增加和分辨率的提升,三维卷积的计算量呈三次方增长,并且此类方法对内存的开销也很大,导致在实际作业场景中无法使用分辨率高的体素网络。而基于点的方法研究的是如何直接处理三维原始点云,主要是为了充分利用点云数据本身的无序性、稀疏性和分布不均等特性,保留信息的完整性。PointNet<sup>[17]</sup>是直接处理点云的开山之作,通过共享的多层感知器(MLP)和空间变换网络学习单点特征,但该方法仅通过最大池化操作聚合所有单点特征得到全局特征,忽略了局部特征信息。之后,PointNet++<sup>[18]</sup>增加的层级结构可以提取点云的局部特征,在对复杂场景和表面纹理丰富的目标进行分类时能取得更好的效果,增加的密度自适应层能够在点云密度不均衡时避免大量冗余计算,提升了网络工作效率。

注意力机制可以极大地提升深度神经网络性能,因此在三维视觉等众多领域得到了广泛应用<sup>[19]</sup>。注意

力机制可以对重要程度不同的信息赋予不同的权重,自适应地挖掘更具鉴别能力的细粒度局部特征<sup>[20]</sup>,从而更加准确地预测。例如:文献[21-22]采用了通道注意力;文献[23]强调信息的具体位置,提出了空间注意力;文献[19]将通道和空间注意力用于PointPillars<sup>[24]</sup>;文献[25]将注意力机制应用于行人检测任务中;文献[26]将注意力机制用于遥感领域;文献[27]在三维空间中采用三重注意力,在加入注意力机制后模型性能得到了显著提升。

在变电站作业场景中采用三维点云数据对目标进行精确定位与检测的工作相对较少,主要难点和挑战体现在以下3个方面:1)变电站内部不仅有众多的电力设备<sup>[28]</sup>,而且电力线路、基座、穿墙套管等附属设施较多<sup>[29-30]</sup>,导致场景复杂度较高;2)电力设备的外廓、尺寸及安装位置均存在较大差异,且即使在变电站内以不同角度采集数据,设备也会受到不同程度的遮挡;3)构造变电站作业场景三维数据集时需要在传感器同步、传感器标定、点云数据标注方面开展大量工作,数据集的缺乏制约了模型的发展。

为了充分利用变电站点云数据的原始特性并提高实际作业场景中目标的检测精度,本文在PointNet++<sup>[18]</sup>中聚合通道方向注意力和点方向注意力,提出了一种适用于变电站作业场景的三维目标检测网络PowerNet。该网络针对原始点云数据进行端到端的三维目标检测,能够更加高效地学习变电站作业场景下点云数据的不同特征信息。在自建的变电站作业场景数据集上的实验结果表明,该网络能够有效地检测变电站场景中变压器设备和行人等目标。

## 2 所提方法

### 2.1 整体网络架构

以PointNet++<sup>[18]</sup>为基础模型,对不同方向的注意力模块进行改进,并将其插入局部特征提取阶段,最终得到了适用于变电站作业场景的三维目标检测网络PowerNet。

PowerNet结构如图1所示。首先对输入的点云数据进行两级局部特征提取,其中每级局部特征提取阶段均包括采样和分组、注意力和mini-pointnet等3个模块。在采样和分组模块(sampling & grouping)<sup>[18]</sup>,采样层对输入点集进行采样,使用迭代最远点采样方法(FPS)选出中心点集,之后分组层利用中心点集以球查询的方法对输入点集进行区域划分,在中心点周围一定半径内取点,得到原始点集的若干局部区域,局部区域的生成过程如图1右下角所示。注意力模块(attention module)分别对上述的若干个局部区域进行处理,对每个区域中的不同信息依据重要程度赋予不同的权重。mini-pointnet模块<sup>[18]</sup>对于每个经注意力模块处理后的局部区域,将每个局部区域内的点的坐标转换为相对该区域中心点的坐标,并将每个局部区域

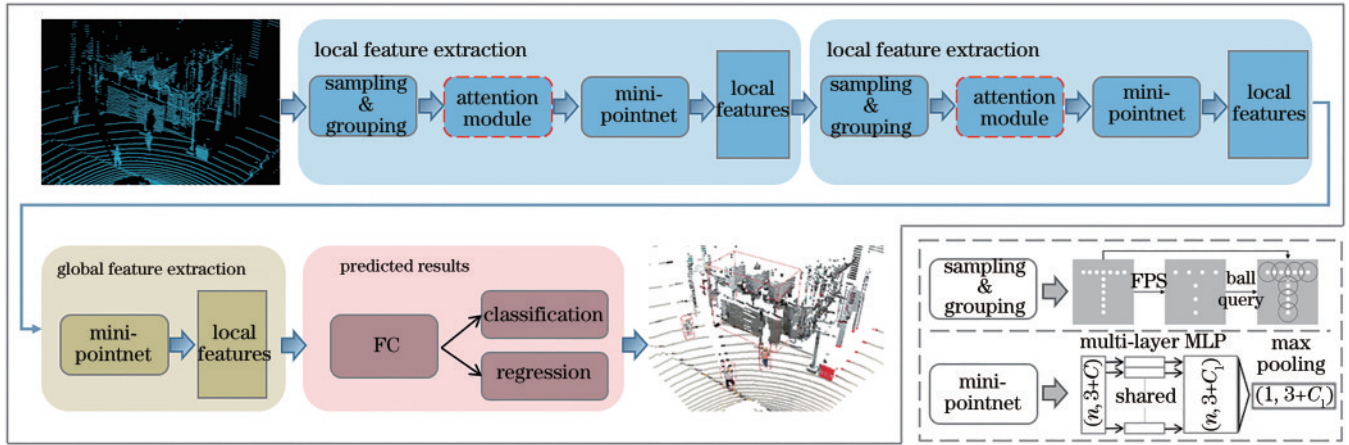


图 1 PowerNet 结构

Fig. 1 PowerNet structure

分别编码为对应的特征向量,得到该区域的局部特征。其次进行全局特征提取,通过 mini-pointnet 模块聚合所有局部特征得到全局特征。mini-pointnet 模块如图 1 右下角所示,其中  $n$  表示点的数量,3 表示点的三维坐标,  $C$  表示每个点的特征维度,  $C_1$  表示经过 MLP 网络升维后点的特征维度。最后由全连接层进行分类与回归,预测目标的类别、中心点坐标、尺寸及角度。

### 2.2 注意力模块

由于点云中不同的点对于检测的重要程度可能不相同<sup>[31]</sup>,目标的特征选择对于识别的效果十分重要<sup>[32]</sup>。为了使网络更加关注重要特征从而提高检测效果,本实验组在局部特征提取阶段插入注意力模块,该模块的输入为分组层得到的局部区域。为充分利用局部区域在通道方向(channel-direction)和点方向(point-direction)的特征信息,本实验组聚合使用了两个方向的注意力。

图 2(a)为该模块输入的球体局部区域,其中  $K$  代表局部区域内点的数量,不同区域  $K$  的大小可能不一样,  $C$  代表每个点编码特征的通道数。为方便对通道方向和点方向注意力的结构图进行描述,将点表示为小立方体,将输入的球体局部区域表征为立方体局部区域,如图 2(b)所示。

#### 2.2.1 通道方向注意力

每个通道都提取特定的特征信息,通道方向注意

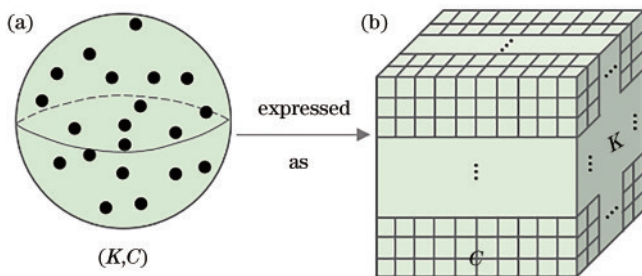


图 2 局部区域。(a)局部区域输入;(b)局部区域表征

Fig. 2 Local area. (a)Local area input; (b) local area representation

力可以建模通道之间的相互依赖。为了提高通道方向注意力的计算效率,同时采用了最大池化和平均池化来聚合点方向上的通道特征。

通道方向注意力结构如图 3 所示。输入一个局部区域  $S_1$  后,分别使用点方向的最大池化和平均池化两种方式得到两个  $1 \times C$  的通道描述,即  $S_{max}^C$  和  $S_{avg}^C$ 。然后将其分别送入一个新的共享网络来为每个通道生成一个权重值。针对变电站环境复杂、设备遮挡等问题,采用更深的 4 层 MLP 网络,并引入残差结构,其中  $r$  为降维比例,  $r$  越小越能保留前一层的全局信息,但计算量会随之增加。为了权衡速度和准确率,将  $r$  取为 16。接着将得到的两个  $1 \times C$  的特征相加后经 Sigmoid 处理得到  $S_1$  的通道方向注意力  $S_c(S_1)$ 。最后将其与输入的局部区域  $S_1$  相乘可得到经过通道方向注意力处理后的局部区域  $S'_1$ 。  $S_1$  的通道方向注意力  $S_c(S_1)$  可以描述为

$$S_c(S_1) = \sigma \left\{ \text{MLP} \left[ \text{MaxPool}(S_1) \right] + \text{MLP} \left[ \text{AvgPool}(S_1) \right] \right\}, \quad (1)$$

$$\text{MLP} \left[ \text{MaxPool}(S_1) \right] = W_3 \left\{ W_2 \left\{ W_1 \left[ W_0(S_{max}^C) \right] \right\} \right\}, \quad (2)$$

$$\text{MLP} \left[ \text{AvgPool}(S_1) \right] = W_3 \left\{ W_2 \left\{ W_1 \left[ W_0(S_{avg}^C) \right] \right\} \right\}, \quad (3)$$

式中:  $S_1$  为输入的局部区域; MaxPool 为最大池化; AvgPool 为平均池化;  $S_{max}^C$  和  $S_{avg}^C$  分别表示经过最大池化和平均池化计算后的局部区域;  $W_0, W_1, W_2, W_3$  分别表示 4 层 MLP 中的参数;  $\sigma$  为 Sigmoid 激活函数。

本实验组采用 4 层 MLP 网络来为每一个特征通道生成权重值,但此时参数之间互相关联的因素会影响模型的精度,可能导致整体梯度的减小甚至消失,深层网络的效果反而变差<sup>[33]</sup>。针对此问题,本实验组在模型中引入残差结构<sup>[34]</sup>,该网络可连接不相邻的网络层,在避免梯度消失问题的同时加深了网络深度,有利于提取更具代表性的深层特征信息。



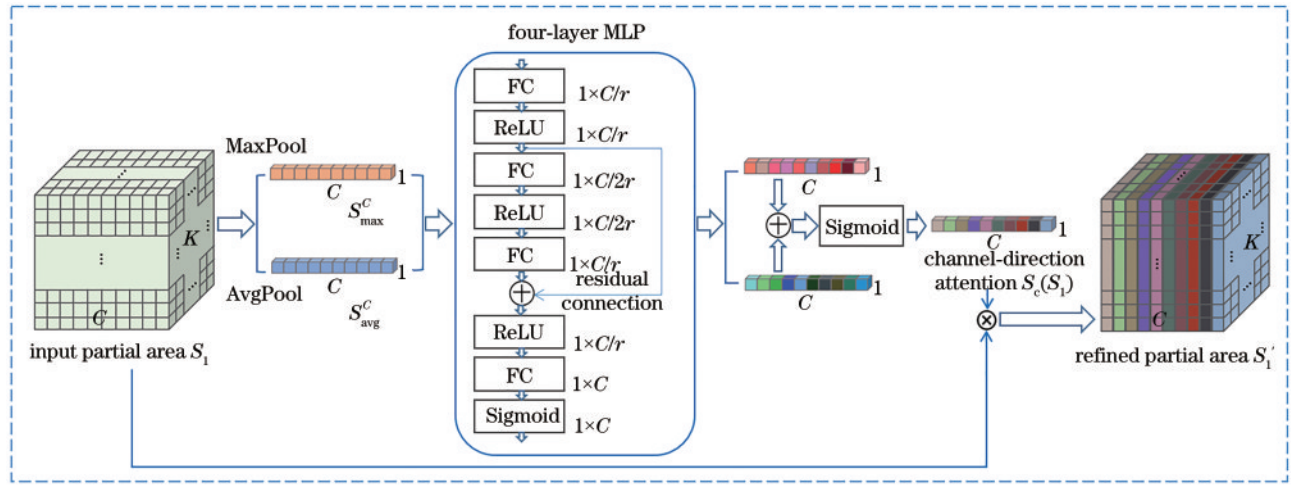


图 3 通道方向注意力结构图  
Fig. 3 Channel direction attention structure diagram

2.2.2 点方向注意力

点方向注意力可以学习局部区域中点之间的相关性,为了在计算点方向注意力时减少信息丢失,在聚合通道方向上的点特征时仍采用最大池化和平均池化两种方式。

点方向注意力结构如图 4 所示。输入一个局部区域  $S_2$  后,分别进行通道维度的最大池化和平均池化,得到两个  $K \times 1$  的点描述,即  $S_{max}^p$  和  $S_{avg}^p$ 。然后将其拼接在一起得到一个通道数为 2 的局部区域,并使用一个包含单个卷积核的隐藏层对其进行卷积操作。在文献[23]的启发下,选择  $5 \times 5$  卷积核,激活函数采用

Sigmoid,卷积时要保证卷积后得到的局部区域在点方向上与输入的局部区域一致,即区域中点的数量不变,卷积后可得到  $S_2$  的点方向注意力  $S_p(S_2)$ 。最后将其与输入的局部区域  $S_2$  相乘可得到经过点方向注意力处理后的局部区域  $S_2'$ 。 $S_2$  的点方向注意力  $S_p(S_2)$  可以描述为

$$S_p(S_2) = \sigma \left\{ f^{5 \times 5} \left[ \text{MaxPool}(S_2), \text{AvgPool}(S_2) \right] \right\}, \quad (4)$$

$$\text{MaxPool}(S_2) = S_{max}^p, \quad (5)$$

$$\text{AvgPool}(S_2) = S_{avg}^p, \quad (6)$$

式中:  $f^{5 \times 5}$  表示使用  $5 \times 5$  的卷积核进行卷积操作。

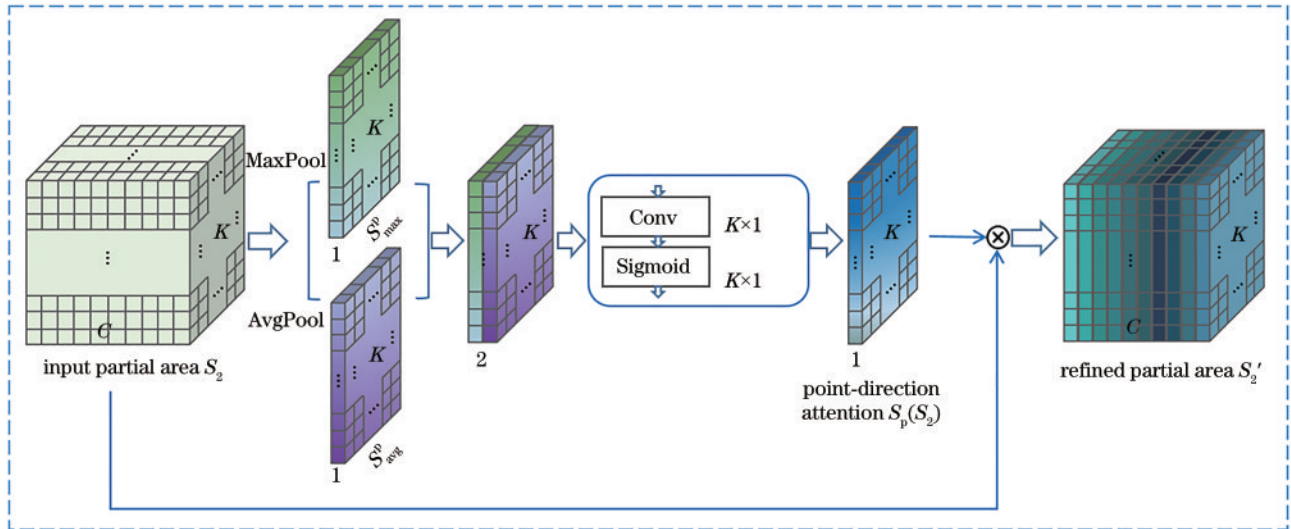


图 4 点方向注意力结构图  
Fig. 4 Point direction attention structure diagram

2.2.3 串行注意力模块

将通道方向注意力与点方向注意力串行连接,用于局部特征提取阶段,如图 5 所示。对于一个局部区域  $\tilde{S}$ ,  $K$  表示该区域内点的数量,不同区域  $K$  的大小可能不一样,  $C$  表示每个点的通道数。首先推理出  $\tilde{S}$  的

通道方向注意力  $S_c(\tilde{S})$ , 将其与输入的局部区域  $\tilde{S}$  相乘,得到经过通道方向注意力加工后的局部区域  $\tilde{S}'$ 。然后计算  $\tilde{S}'$  的点方向注意力  $S_p(\tilde{S}')$ , 将其与  $\tilde{S}'$  相乘得到经过通道和点方向双重注意力串行加工后的局部区

域  $\tilde{S}''$ 。最后将  $\tilde{S}$  和  $\tilde{S}''$  的加权和作为输出的局部区域  $\tilde{S}'''$ 。具体计算过程如下：

$$\tilde{S}' = S_c(\tilde{S}) \otimes \tilde{S}, \quad (7)$$

$$\tilde{S}'' = S_p(\tilde{S}') \otimes \tilde{S}', \quad (8)$$

$$\tilde{S}''' = \tilde{S} \oplus \tilde{S}'' \quad (9)$$

式中： $\otimes$ 代表逐元素相乘； $\oplus$ 代表加权求和。

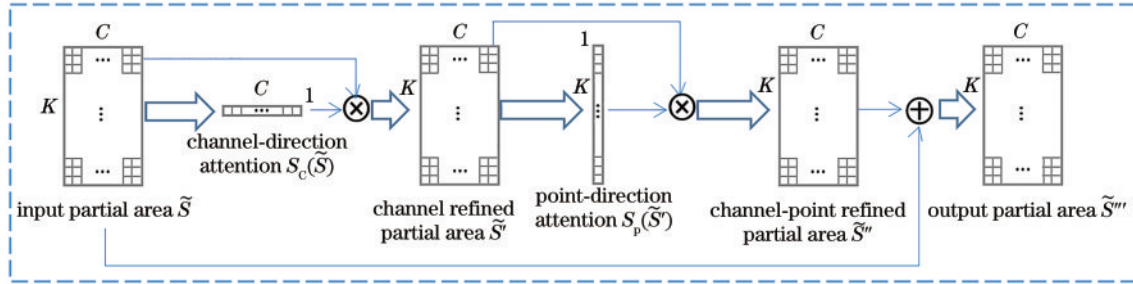


图 5 串行注意力结构图

Fig. 5 Serial attention structure diagram

### 2.3 损失函数

针对变电站点云数据中目标前景点与背景点的数量差距较大,普通的交叉熵损失函数在挖掘前景点的重要性方面表现不够好的问题,本实验组将基础模型 PointNet++<sup>[18]</sup>中的交叉熵损失替换为 Focal 损失<sup>[35]</sup>来计算分类损失,使训练更加关注前景点,角度损失及回归损失的计算沿用文献[27]的方法。Focal 损失的表达式为

$$L_{cls} = \begin{cases} -\alpha(1-p)^{\gamma} \ln p, & y = 1 \\ -(1-\alpha)p^{\gamma} \ln(1-p), & y \neq 1 \end{cases}, \quad (10)$$

式中： $p$ 为预测概率； $\alpha$ 为权重因子,用来平衡正负样本的重要性； $\gamma$ 为调制因子,用来调节样本权重降低的速率； $y$ 为真值的类别, $y = 1$ 时预测概率为  $p$ 。 $\alpha$ 取 0.25, $\gamma$ 取 2。

## 3 实验

### 3.1 实验环境及评价指标

采用 Ubuntu 18.04 64 位操作系统、NVIDIA

GeForce RTX 2080Ti 显卡、Intel i9-9900K CPU 进行实验,CUDA 采用 10.1 版本,CUDNN 采用 7.6.5 版本,Python 采用 3.6 版本,网络基于 Pytorch 1.1.0 深度学习框架实现。

训练过程选用 SGD 优化器,初始学习率为 0.1,采用指数衰减方式,权重衰减系数为 0.2,动量为 0.9,训练时仅输入点云数据,共训练 1000 个 epoch,batch size 设置为 4。将变压器类别的交并比 (IoU) 值设置为 0.7,行人类别的 IoU 设置为 0.5。

采用均值平均精度 (mAP) 作为评价指标,mAP 表示不同类别的平均精度 (AP) 均值。

### 3.2 数据集

#### 3.2.1 数据采集

采用 Ouster OS-1-64 线激光雷达采集点云数据,采用 120° 广角摄像头采集 RGB 图像,采集器如图 6(a) 所示。所用数据集为某 220 kV 变电站作业场景下的点云和 RGB 图像,规模为 1792,包含两个类别,分别是行人和变压器,如图 6(b) 所示。训练集、验证集、测试集按 6:2:2 的比例进行划分。

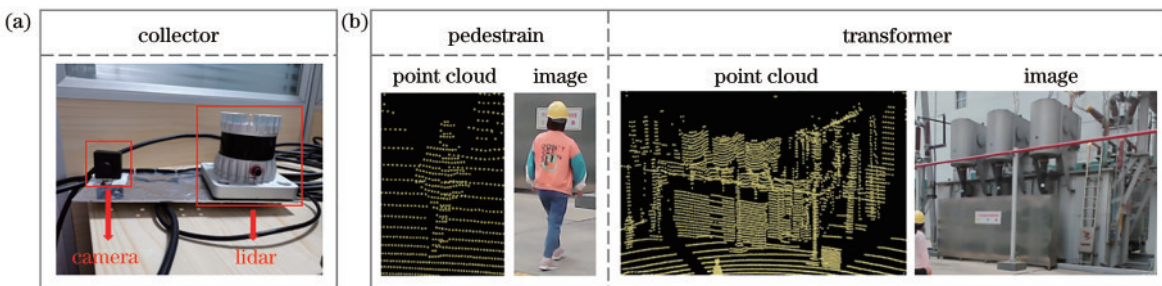


图 6 数据集采集器及样本图示。(a)数据集采集器;(b)样本图示

Fig. 6 Dataset collector and sample illustration. (a) Dataset collector; (b) sample illustration

#### 3.2.2 数据标注

分别采用 PCAT 点云标注工具和 LabelImg 图像标注工具对采集到的点云和 RGB 图像中的目标进行标注,如图 7(a)、(b) 所示。按照 KITTI 数据集标签格

式对两者进行整理,得到的标签格式如图 7(c) 所示,共包括 15 个属性值,分别为物体类别、是否截断、是否遮挡、观察角度、二维物体边界框、三维物体尺寸(高、宽、长)、三维物体位置、三维物体全局方向角。

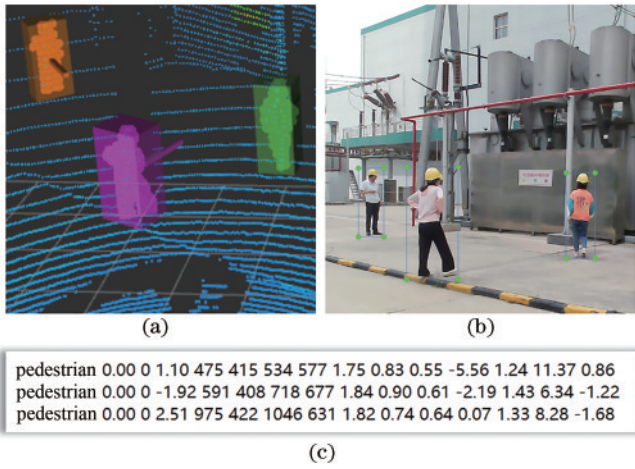


图 7 数据标注。(a)PCAT 标注点云;(b)Labellmg 标注图像;(c)标签格式

Fig. 7 Data annotation. (a) PCAT annotated point cloud; (b) Labellmg annotated images; (c) label format

### 3.3 实验结果及分析

#### 3.3.1 注意力的不同组合对网络性能影响的对比实验

为分析不同因素对网络性能的影响,在获得通道

方向注意力时采用不同深度的 MLP,在获得点方向注意力时采用不同大小的卷积核,并对通道方向注意力和点方向注意力采用并行(parallel)和串行(serial)两种连接方式进行对比实验,结果如表 1 所示,粗体表示最优结果。

从表 1 可以看出,在获得通道方向注意力时采用更深层次的 4 层 MLP 网络,在获得点方向注意力时采用大小为  $5 \times 5$  的卷积核,注意力的连接模式为串行时,可以取得更好的检测效果,mAP 值达到 0.735。因此,所提方法采用 4 层 MLP 网络和大小为  $5 \times 5$  的卷积核,并将两种注意力串行连接。

#### 3.3.2 注意力结构的选择

为了对比不同的注意力结构对检测精度的影响,对通道方向注意力和点方向注意力进行了实验,实验结果如表 2 所示,其中  $\checkmark$  表示使用相应方法,-表示未使用相应方法。从表 2 可以看出,分别使用通道方向注意力结构和点方向注意力结构进行实验,与不采用注意力结构相比,mAP 值分别提升了 2.1 个百分点和 12.1 个百分点,即通道方向注意力结构更有利于提升模型检测精度。将两种注意力串行后,mAP 值达到 0.735,验证了所提方法的实用性。

表 1 注意力的不同组合对网络性能影响的对比结果

Table 1 Comparison of effects of different combinations of attention on network performance

Channel-direction attention	Point-direction attention	Parallel			Serial		
		AP		mAP	AP		mAP
		Pedestrian	Transformer		Pedestrian	Transformer	
Two-layer MLP	$7 \times 7$ filter	0.550	0.776	0.663	0.572	0.797	0.685
Two-layer MLP	$5 \times 5$ filter	0.559	0.781	0.670	0.576	0.800	0.688
Four-layer MLP	$7 \times 7$ filter	0.572	0.794	0.683	0.591	0.849	0.720
Four-layer MLP	$5 \times 5$ filter	0.579	0.802	0.691	<b>0.602</b>	<b>0.867</b>	<b>0.735</b>

表 2 注意力结构的选择

Table 2 Choice of attention structure

Channel-direction attention (four-layer MLP)	Point-direction attention ( $5 \times 5$ filter)	AP		mAP
		Pedestrian	Transformer	
-	-	0.545	0.775	0.660
$\checkmark$	-	0.572	0.790	0.681
-	$\checkmark$	0.560	0.779	0.670
$\checkmark$	$\checkmark$	<b>0.602</b>	<b>0.867</b>	<b>0.735</b>

#### 3.3.3 损失函数的选择

为了分析不同的分类损失函数对模型检测精度的影响,分别将交叉熵损失和焦点损失作为分类损失函数进行实验,实验结果如表 3 所示。从表 3 可以看出,采用焦点损失替换交叉熵损失后,mAP 值提升了 1.5 个百分点。即针对变电站点云数据中目标前景点与背景点数量不平衡的情况,采用焦点损失函数计算分类损失是有效的。

表 3 损失函数的选择

Table 3 Choice of loss function

Cross entropy loss	Focal loss	AP		mAP
		Pedestrian	Transformer	
$\checkmark$	-	0.572	<b>0.868</b>	0.720
-	$\checkmark$	<b>0.602</b>	0.867	<b>0.735</b>



## 3.3.4 与主流检测模型的性能对比实验

与目前主流的三维检测模型进行了对比实验,结果如表 4 所示。所提方法对行人和变压器的 AP

值分别达到 0.602 和 0.867, mAP 值达到 0.735, 检测精度超越其他方法。即所提方法具有较高的检测性能。

表 4 主流检测模型性能对比结果

Table 4 Performance comparison results of mainstream detection models

Method	Model	AP		mAP
		Pedestrian	Transformer	
3D to 2D	PIXOR <sup>[9]</sup>	0.527	0.755	0.641
	Complex-YOLO <sup>[10]</sup>	0.533	0.779	0.656
Voxelization	Vote3Deep <sup>[15]</sup>	0.537	0.733	0.635
	VoxelNet <sup>[13]</sup>	0.531	0.802	0.667
Original point cloud	PointNet <sup>[17]</sup>	0.540	0.762	0.651
	PointNet++ <sup>[18]</sup>	0.545	0.775	0.660
	Proposed method	<b>0.602</b>	<b>0.867</b>	<b>0.735</b>

## 3.3.5 检测结果可视化

所提方法训练过程中损失变化情况及不同类别的 AP 变化情况如图 8 所示,横轴表示训练轮数,纵轴表示对应数值。损失曲线表示损失变化情况,随着训练次数的增加,损失不断下降,模型逐渐收敛。AP1 和 AP2 曲线分别代表行人和变压器的 AP 变化情况,随着训练的进行,均稳步提升,最终分别稳定在 0.602 和 0.867。所提方法的检测结果图如图 9 所示,其中图 9(a)为 RGB 图像,图 9(b)为对应的点云检测结果图。可视化结果表明,所提方法在复杂的变电站环境下能够较好地检测行人和变压器等 3D 目标,进一步验证了其有效性。

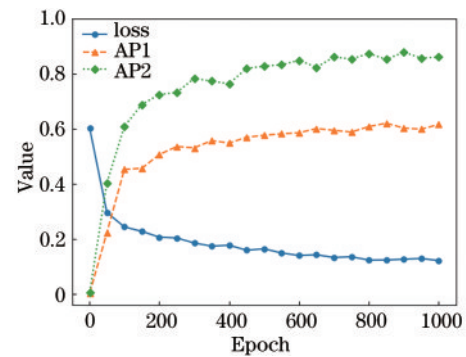


图 8 损失曲线和性能曲线

Fig. 8 Loss curve and performance curve

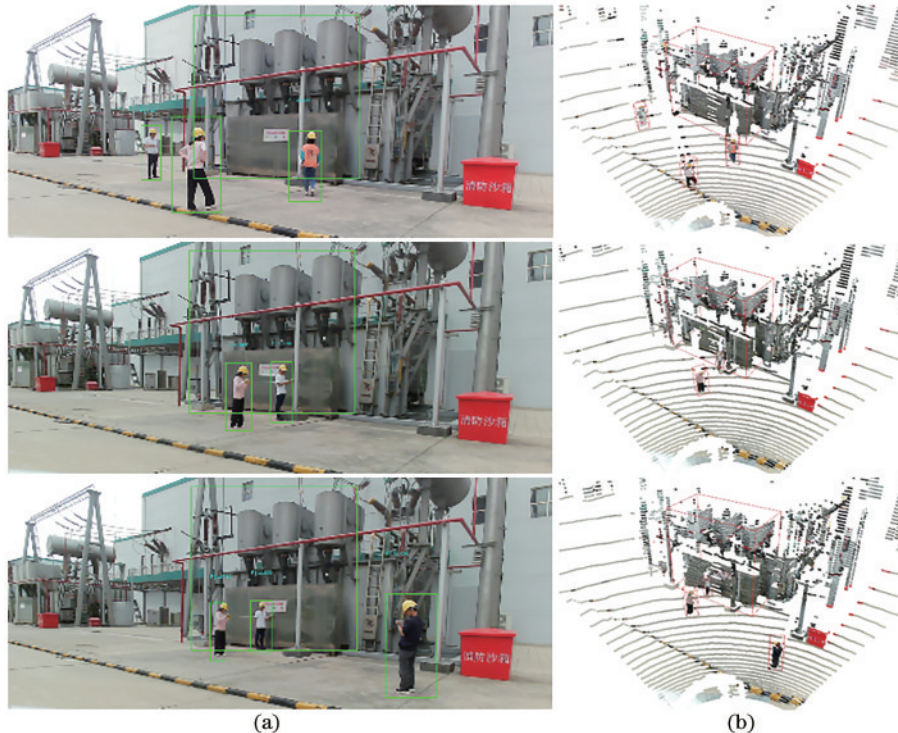


图 9 检测结果。(a)RGB 图像;(b)点云图

Fig. 9 Test result. (a) RGB images; (b) point cloud diagrams

## 4 结 论

针对变电站作业场景中的安全距离管控工作依赖三维目标的精确定位,且变电站中存在背景复杂、设备遮挡等问题,提出了一种在变电站环境中基于注意力机制的变电站作业场景 3D 目标检测模型 PowerNet。实验结果表明,将通道方向注意力和点方向注意力串行连接后,PowerNet 的性能优于其他网络。该网络对于变电站作业场景下的目标检测具有一定的实用价值。接下来,将进一步优化模型性能,在下游安全距离管控任务中开展应用测试。

### 参 考 文 献

- [1] 宁日红. 变电站运行状态监测及其风险评估[J]. 科技风, 2021, (19): 190-191.  
Ning R H. Operation status monitoring and risk assessment of substations[J]. Technology Wind, 2021, (19): 190-191.
- [2] 刘彭瑀. 智能变电站继电保护的运行和维护管理[J]. 技术与市场, 2021, 28(7): 179-180.  
Liu P Y. Operation and maintenance management of intelligent substation relay protection[J]. Technology and Market, 2021, 28(7): 179-180.
- [3] 金学煜. 变电站安全作业移动定位与报警终端研究[J]. 信息与电脑(理论版), 2020, 32(19): 97-99.  
Jin X Y. Research on mobile location and alarm terminal for safe operation in substation[J]. China Computer & Communication, 2020, 32(19): 97-99.
- [4] 罗克伟. 基于三维数字化平台的变电站智能可视化运维系统研究[J]. 电工技术, 2020, 4(11): 155-158.  
Luo K W. Research on intelligent visualization management system for substation base on three-digital platform[J]. Electric Engineering, 2020, 4(11): 155-158.
- [5] 孙海渤, 岳红原, 王智睿, 等. 变电站三维可视化管控系统的设计与实现[J]. 工业控制计算机, 2020, 33(2): 97-99.  
Sun H B, Yue H Y, Wang Z R, et al. 3D visualization management-control system for substation[J]. Industrial Control Computer, 2020, 33(2): 97-99.
- [6] 彭磊. 三维激光扫描仪在变电站改扩建工程中的应用[J]. 河北电力技术, 2020, 39(5): 40-42.  
Peng L. Application of 3D laser scanner in substation reconstruction and expansion project[J]. Hebei Electric Power, 2020, 39(5): 40-42.
- [7] 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.
- [8] 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.
- [9] Yang B, Luo W J, Urtasun R. PIXOR: real-time 3D object detection from point clouds[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7652-7660.
- [10] Simon M, Milz S, Amende K, et al. Complex-YOLO: an Euler-region-proposal for real-time 3D object detection on point clouds[M]//Leal-Taixé L, Roth S. Computer vision-ECCV 2018 workshops. Lecture notes in computer science. Cham: Springer, 2019, 11129: 197-209.
- [11] Beltran J, Guindel C, Moreno F M, et al. BirdNet: a 3D object detection framework from LiDAR information[C]//2018 21st International Conference on Intelligent Transportation Systems (ITSC), November 4-7, 2018, Maui, HI. New York: IEEE Press, 2018: 3517-3523.
- [12] Chen X Z, Ma H M, Wan J, et al. Multi-view 3D object detection network for autonomous driving[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI. New York: IEEE Press, 2017: 1907-1915.
- [13] Zhou Y, Tuzel O. VoxelNet: end-to-end learning for point cloud based 3D object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4490-4499.
- [14] Li B. 3D fully convolutional network for vehicle detection in point cloud[C]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 24-28, 2017, Vancouver, BC, Canada. New York: IEEE Press, 2017: 1513-1518.
- [15] Engelcke M, Rao D, Wang D Z, et al. Vote3Deep: Fast object detection in 3D point clouds using efficient convolutional neural networks[C]//2017 IEEE International Conference on Robotics and Automation, May 29-June 3, 2017, Singapore. New York: IEEE Press, 2017: 1355-1361.
- [16] Shi S S, Wang Z, Shi J P, et al. From points to parts: 3D object detection from point cloud with part-aware and part-aggregation network[EB/OL]. (2017-06-07) [2021-02-05]. <https://arxiv.org/abs/1907.03670>.
- [17] Charles R Q, Hao S, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [18] Charles R Q, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [EB/OL]. (2017-06-07) [2021-02-05]. <https://arxiv.org/abs/1706.02413>.
- [19] 詹为钦, 倪蓉蓉, 杨彪. 基于注意力机制的 PointPillars+ 三维目标检测[J]. 江苏大学学报(自然科学版), 2020, 41(3): 268-273.  
Zhan W Q, Ni R R, Yang B. An attention-based PointPillars+ 3D object detection[J]. Journal of Jiangsu University (Natural Science Edition), 2020, 41(3): 268-273.
- [20] 于丽丽, 于海洋, 何子鑫, 等. 基于双注意力机制和多尺度特征的点云场景分割[J]. 激光与光电子学进展, 2021, 58(24): 2428007.



- Yu L L, Yu H Y, He Z X, et al. Point cloud scene segmentation based on dual attention mechanism and multi-scale features[J]. *Laser & Optoelectronics Progress*, 2021, 58(24): 2428007.
- [21] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [22] Zagoruyko S, Komodakis N. Paying more attention to attention: improving the performance of convolutional neural networks via attention transfer[EB/OL]. (2016-12-12)[2021-02-05]. <https://arxiv.org/abs/1612.03928>.
- [23] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [24] Lang A H, Vora S, Caesar H, et al. PointPillars: fast encoders for object detection from point clouds[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: 12689-12697.
- [25] 邹梓吟, 盖绍彦, 达飞鹏, 等. 基于注意力机制的遮挡行人检测算法[J]. *光学学报*, 2021, 41(15): 1515001.
- Zou Z Y, Gai S Y, Da F P, et al. Occluded pedestrian detection algorithm based on attention mechanism[J]. *Acta Optica Sinica*, 2021, 41(15): 1515001.
- [26] 田青林, 秦凯, 陈俊, 等. 基于注意力金字塔网络的航空影像建筑物变化检测[J]. *光学学报*, 2020, 40(21): 2110002.
- Tian Q L, Qin K, Chen J, et al. Building change detection for aerial images based on attention pyramid network[J]. *Acta Optica Sinica*, 2020, 40(21): 2110002.
- [27] Liu Z, Zhao X, Huang T T, et al. TANet: robust 3D object detection from point clouds with triple attention[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 11677-11684.
- [28] 龙丽娟, 夏永华, 黄德. 一种基于三维激光扫描点云数据的变电站快速建模方法[J]. *激光与光电子学进展*, 2020, 57(20): 202801.
- Long L J, Xia Y H, Huang D. Fast modeling method for substation based on 3D laser scanning point cloud data[J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 202801.
- [29] 王菲, 王球, 任佳依, 等. 变电站电气设备检测与三维建模系统[J]. *电测与仪表*, 2021, 58(3): 160-167.
- Wang F, Wang Q, Ren J Y, et al. Substation electrical equipment detection and three-dimensional modeling system[J]. *Electrical Measurement & Instrumentation*, 2021, 58(3): 160-167.
- [30] 王菲, 王球, 任佳依, 等. 三维激光云扫描技术在变电站电气设备识别中的应用[J]. *佳木斯大学学报(自然科学版)*, 2020, 38(4): 124-128.
- Wang F, Wang Q, Ren J Y, et al. Application of 3D laser cloud scanning technology in identification of electrical equipment in substation[J]. *Journal of Jiamusi University (Natural Science Edition)*, 2020, 38(4): 124-128.
- [31] 万思宇. 基于注意力机制的3D车辆检测算法[J]. *计算机工程与科学*, 2020, 42(1): 98-102.
- Wan S Y. A 3D vehicle detection algorithm based on attention mechanism[J]. *Computer Engineering & Science*, 2020, 42(1): 98-102.
- [32] 刘丹丹. 基于表面特征的变电站设备三维点云识别[D]. 郑州: 郑州大学, 2016: 1-6.
- Liu D D. 3D point cloud of substation equipment recognition based on surface features[D]. Zhengzhou: Zhengzhou University, 2016: 1-6.
- [33] 陈文豪, 何敬, 刘刚. 引入注意力机制的卷积神经网络高光谱图像分类[J]. *激光与光电子学进展*, 2022, 59(18): 1811001.
- Chen W H, He J, Liu G. Hyperspectral image classification based on convolution neural network with attention mechanism[J]. *Laser & Optoelectronics Progress*, 2022, 59(18): 1811001.
- [34] 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 (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [35] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.