doi: 10. 3969 / j. issn. 1674 - 1242. 2023. 04. 011

# 基于多尺度特征增强的小样本计数

## 郑东宏, 何志权

(广东省多媒体信息服务工程技术研究中心,深圳大学电子与信息工程学院,广东深圳 518060)

【摘要】目标物体计数是计算机视觉领域的重要研究方向。针对小样本计数中存在的样本与查询图像目标物体尺寸不一致、目标物体分布不均匀的问题,该文提出了多尺度特征增强计数算法。首先,基于特征金字塔构建自上而下的特征融合网络。在各级尺度上对查询图像中和样本相似度较高的区域进行样本特征增强,随后送入上一级特征匹配。然后,将各级增强后的查询特征送入回归器中,得到各级密度图。最后,求和,生成高质量的密度图。该文在FSC-147和 CARPK 数据集上进行测试。实验结果表明,该文所提模型的性能优于大多数其他方法,有效改善了目标物体大小变化造成的计数精度低的问题。

【关键词】小样本; 计数; 多尺度; 特征增强

【中图分类号】TP18

【文献标志码】A

文章编号: 1674-1242 (2023) 04-0405-07

## Multi-scale Feature Enhancement for Few-shot Object Counting

ZHENG Donghong, HE Zhiquan

(Guangdong Engineering Research Center of Multimedia Information Service, School of Electronic and Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China)

**[Abstract]** Object counting is an important research direction of computer vision. In order to solve the problem that the size of the object in the few-shot object counting is inconsistent with that in the query image, and the object distribution is uneven, a multi-scale feature enhancement counting algorithm is proposed. First, a top-down feature fusion network is constructed based on the feature pyramid. At each level of scale, the sample features of the regions with high sample similarity in the query image are enhanced, and then sent to the upper level for feature matching. The enhanced query features at all levels are sent into the regression head to obtain the density maps at all levels, and finally the sum can generate a high-quality density map. The experimental results on FSC-147 and CARPK datasets show that the performance of the model is better than that of most previous methods, which effectively improves the problem of low counting accuracy caused by the change of target size.

**[Key words]** Few-shot; Counting; Multi-scale; Feature Enhancement

## 0 引言

对图像中的特定目标物体进行计数是计算机视觉 领域的一项重要任务,在人群计数、车辆计数、农产 品计数等许多领域均有广泛的应用。小样本计数是指 给定1至多个样本,在查询图像中寻找与样本类别相同的目标物体并计数。该方法不要求目标物体是在训练数据中出现过的类别,用户可以在图像中任意指定多个样本,将样本和查询图像输入模型中,模型输出

收稿日期: 2022-12-04。

基金项目:国家自然科学基金(61971290)。

作者简介:郑东宏(1998-),男,广东省汕头市人,硕士研究生,从事物体计数研究。

通信作者:何志权,男,副教授,硕士研究生导师,电话(Tel.):18680358296,邮箱(E-mail):zhiquan@szu.edu.en。

查询图像对应的样本类别物体的计数值。如此即可用 一个模型来完成不特定类别物体的计数。本文提出了 一种新颖的多尺度特征增强计数算法。本文的贡献如 下:(1)构建了多尺度网络来捕捉不同层次的语义信 息,遵循卷积神经网络的多尺度和分层结构,以由粗 到细的方式进行特征匹配,在不同的尺度上产生互补 的相似图。(2)提出了样本特征增强模块,对查询图 像中和样本相似度较高的区域进行样本特征增强,将 增强后的查询图像送入回归器中,得到计数值。

# 1 计数方法

#### 1.1 传统计数方法

在计数领域,传统方法是指使用传统的计算机视 觉和图像处理技术进行物体计数。这些方法通常基于 图像的像素级别处理和特征提取,使用算法进行计数。

以下是一些常见的传统计数方法。

(1)阈值分割法。该方法基于图像的颜色和亮度 信息,首先通过设定一个阈值将前景物体与背景分离, 然后通过计算前景区域的像素数进行计数。这种方法 简单直接,但对于复杂的场景和变化的光照条件可能 欠缺鲁棒性。

(2)边缘检测法。该方法首先使用边缘检测算法 (如 Canny 算子)来检测图像中的边缘。然后通过对边 缘进行连接和处理,得到物体的轮廓。最后通过计算 轮廓的数量或连接相邻轮廓进行计数。这种方法对于 清晰的边缘和简单的场景比较有效,但对于复杂的形 状和遮挡情况可能不够准确。

(3)区域分割法。该方法将图像分割成多个区域, 对每个区域进行特征提取和分析。常用的区域分割法 包括基于颜色、纹理或形状的分割方法。通过对每个 区域进行特征提取和计数,可以得到总体的物体计数。 这种方法对于复杂的场景和多个物体之间的遮挡有一 定的鲁棒性。

(4)模板匹配法。该方法使用预定义的物体模板 进行匹配来计数。首先,提前准备好代表物体的模板 图像。然后,将模板与输入图像进行匹配,找到与模 板相似的物体。通过计数匹配到的物体数量,可以进 行物体计数。这种方法对于形状和尺寸相对固定的物 体比较有效,但对于形状变化较大的物体可能不太 适用。

#### 1.2 基于深度学习的特定类别物体计数

通过对特定类别目标物体的计数,可以从视频或

图像中得到物体的空间分布信息,为进一步的数据分 析奠定基础。以人群计数<sup>[1]</sup>为例,在游行集会中,通 过对人群密度的分析,可以及时疏散人群,避免踩踏 事故的发生。车辆计数<sup>[2]</sup>则有助于城市交通指挥系统 实时了解道路拥堵情况,缓解交通压力。

随着深度学习的快速发展,基于深度学习的应用 越来越多<sup>[3,4]</sup>。在特定类别物体的计数上,采用深度学 习方法的计数效果与传统计数方法相比,有了极大的 提高。卷积神经网络凭借其对图像深度特征出色的提 取能力而被广泛应用于目标计数中。基于深度学习的 计数方法主要划分为基于回归的检测计数方法和基于 密度图的估计方法。前者依赖目标检测器,通过目标 检测得到物体的定位,同时可以统计目标物体数目。 近年来,目标检测算法快速发展,YOLO<sup>[5]</sup>、 RetinaNet<sup>[6]</sup>、CenterNet<sup>[7]</sup>等计数方法在目标分类和目 标定位上的准确度不断提高。但目标检测算法并非专 门为计数领域设计的,需要训练针对不同类别物体的 检测器,且训练时需要的注释信息远比基于密度图估 计方法的多。此外,当物体较密集,存在遮挡、重叠 等现象时,其性能并不能令人满意。基于密度图的估 计方法[8-10]从查询图像中提取图像特征,利用该特征预 测每个像素的密度值。其在训练时只需少量注释信息, 注释数据集时只需在每个目标物体中心标注"1",用 高斯核卷积查询图像的真实计数图,即可生成密度图。 与矩形框注释相比,这种类型的注释效率更高,劳动 强度更小。在基于密度图的估计方法中, Zeng 等<sup>[11]</sup> 提出了基于多尺度人群特征融合计数的方法,提取人 群在不同尺度上的特征并进行特征融合计数。Wang 等[12]以膨胀卷积理论为基础,构建多尺度特征提取模 块,以适应密集人群在尺度上分布不均匀的特性。上 述方法都是针对特定类别物体的计数方法。

#### 1.3 小样本计数

小样本学习是为了解决在训练数据集不足的情况 下,深度学习模型难以充分学习和泛化的问题。在许 多实际场景中,由于数据的获取成本高、标注困难或 样本稀缺,可能只有很少的标注样本可用于模型训练。 这导致传统的深度学习方法难以获得良好的性能。

小样本学习旨在通过设计新的学习算法和策略, 使深度学习模型能够在少量样本下进行有效学习,以 解决以下问题。 (1)过拟合。在少量样本情况下,深度学习模型 很容易过拟合,即在训练集上表现良好,但在新数据 上表现不佳。小样本学习通过引入正则化技术、模型 复杂度控制和参数约束等方法,可以减少过拟合风险。

(2)数据不平衡。在少量样本情况下,不同类别 或不同场景之间的数据分布可能不平衡,即某些类别 或场景的样本数量较少。小样本学习可以通过样本加 权、重采样和数据增强等技术平衡数据分布,从而提 高模型在少数类别物体或场景中的性能。

(3)迁移学习。在少量样本情况下,传统的从头 开始训练深度学习模型往往无法收敛或得到良好的结 果。小样本学习中的迁移学习方法可以利用预训练模 型在大规模数据中学到的通用特征表示,加快模型在 小样本任务中的学习速度和性能。

(4)元学习。元学习是一种能够从少量样本中快速学习和适应新任务的学习方法。小样本学习中的元 学习可以帮助模型从少量样本中快速捕捉到任务的特 定特征和模式,从而提高模型在小样本任务中的性能。

传统的深度学习计数方法需要大量的训练数据来 学习针对特定类别物体的计数模型,当目标物体的种 类发生变化时,计数效果往往不尽如人意,这极大地 限制了计数模型的应用。此外,受成本和注释的难度 等因素的影响,收集大量的同类别物体标记数据并不 容易,特别是在一些专业领域,如医学领域和生物科 学领域。训练模型所消耗的时间和资金也是限制传统 深度学习计数模型应用的重要影响因素。为了解决计 数模型的泛化问题,研究者<sup>[13]</sup>结合小样本学习提出了 小样本计数(Few-Shot Object Counting, FSC),又称 为类未知计数(Class Agnostic Counting, CAC)。

小样本计数将小样本学习的元学习思想和深度学 习计数方法相结合,主要包括特征提取和特征匹配两 个阶段。首先从样本中提取视觉特征,并将这些特征 与查询图像的特征进行匹配,以相似度匹配结果作为 推断物体数量的重要依据。GMN<sup>[11]</sup>对样本特征进行池 化,并将池化结果连接到查询图像特征上,然后送入 回归器中产生密度图。FamNet<sup>[14]</sup>通过多尺度提高相似 图的可靠性,用相似图进行计数,但是相似图远不如 特征图的信息量大,预测密度图边界较模糊。

#### 2 多尺度特征增强网络结构设计

2.1 多尺度网络

多尺度网络是一种深度神经网络结构,用于处理

具有不同尺度信息的图像数据。它的设计灵感来自人 类视觉系统的特点,即人类在分析和理解图像时会同 时关注不同尺度的视觉特征。

多尺度网络通常由多尺度输入、特征提取模块、 尺度融合、多尺度输出构成。多尺度网络接受输入图 像的不同尺度版本。这可以通过对原始图像进行金字 塔尺度变换、图像金字塔或多尺度裁剪等方式实现。 每个尺度的图像输入将包含不同的细节和上下文信 息。多尺度网络包含多个并行的特征提取模块,每个 模块专门处理一个特定尺度的输入。这些模块可以是 卷积神经网络或其他特征提取器,用于从不同尺度的 图像中提取特征表示。模块之间的参数一般是共享的, 以提高模型的效率和泛化能力。

多尺度网络通过将不同尺度提取的特征进行融合,综合各个尺度的信息。融合可以通过卷积、池化、 拼接或注意力机制等方式进行,目的是使多尺度网络 能够同时利用多尺度信息来更全面地理解图像。多尺 度网络的输出可以包括多个尺度的预测结果。这些结 果可以是分类、定位、分割或其他任务的输出。通过 在不同尺度上生成输出,多尺度网络可以更好地处理 不同尺度的目标物体或图像细节。

多尺度网络的优势在于它能够处理具有不同尺度 信息的图像,从而提高对目标物体的感知能力和对细 节的理解能力。通过利用多尺度信息,网络可以更好 地适应不同尺度的物体、背景噪声和环境变化。多尺 度网络在许多计算机视觉任务中具有广泛的应用,如 目标检测、图像分割、姿态估计等,可以提高模型的 性能和鲁棒性。

在深度特征提取网络中,接近输入图像的底层卷 积层对颜色、边缘、纹理等低层信息非常敏感,相关 的卷积核的感受野小,缺乏上下文信息,获得的局部 信息多,能检测到小物体。顶层卷积层具有丰富的高 级语义,其感受野大,对大物体检测效果好,但容易 漏检小目标。多尺度特征提取网络将各个尺度下的语 义信息结合起来,融合各层语义信息,实现优势互补。

本文以 Resnet 50<sup>[15]</sup>作为特征提取器,获得在各个 尺度下 Resnet 50 残差块输出的特征。采用 U-Net<sup>[16]</sup> 通道串联结构作为主干网络,自下而上将样本特征增 强后的查询图像特征逐层向上融合,以由粗到细的方 式实现多尺度信息融合。多尺度特征增强网络结构示 意如图 1 所示。



Fig. 1 Illustration of multi scale feature enhancement network

#### 2.2 特征增强

多个样本的特征图长宽不一,为了便于模型训练 计算,将N个样本特征送入自适应最大值池化层,设 置样本特征图尺寸为5×5,通道数与查询特征图一致。 以N个5×5样本特征图作为卷积核,将其与查询特征 图 F<sub>a</sub>进行卷积,得到N张相似图S。

$$S = F_Q * F_{S(\text{kernel})} \tag{1}$$

对每幅相似图 S<sub>i</sub>做归一化处理。

$$S'_{i} = \frac{e^{s_{i}}}{\max(e^{s_{i}})} \tag{2}$$

相似图的大小与查询特征图一致,通道数为 1。 以相同的 N 个 5×5 样本特征图作为卷积核,将其与对 应的相似图进行卷积,得到 N 个相似图指引的样本增 强特征图,对其求和得到 F'o。

$$F_Q' = \sum S' \times F_{S(\text{kernel})} \tag{3}$$

*F'<sub>Q</sub>*的维度大小与*F<sub>Q</sub>*相同。将*F'<sub>Q</sub>*通过卷积层与*F<sub>Q</sub>*相加,即可得到相应尺度下增强样本特征后的查询图像特征图*F*"。

$$F_O'' = \operatorname{Conv}(F_O') + F_O \tag{4}$$

#### 2.3 多尺度增强特征图计数

将 F"进行层归一化处理,通过卷积层上采样到上一级尺度中进行通道拼接,同时在原先的尺度上实现特征增强,之后送入回归器中进行计数。回归器由卷积层和上采样层构成,在卷积层中将通道数逐步降低到1,通过上采样恢复成原始查询图像大小,回归器

输出的张量通道数为1,长和宽与原始查询图像一致。 将 Resnet 50 三个尺度所得到的密度图进行求和,即可 得到密度图。

## 2.4 损失函数

损失函数(Loss Function)是在机器学习和优化 问题中使用的一种函数,用于衡量模型的预测输出与 实际目标之间的差异或损失程度。损失函数在训练过 程中起着重要的作用,它提供了一个可优化的目标, 使模型能够根据实际情况不断调整自身的参数,以最 小化损失函数的值。本文采用L2损失作为损失函数。

$$L = \left\| D_{\text{pre}} - D_{\text{gt}} \right\|_{2}^{2} \tag{5}$$

式中, D<sub>pre</sub> 是模型输出的预测密度图, D<sub>gt</sub> 是真实密度图。

#### 3 实验

#### 3.1 数据集准备

FSC-147<sup>[14]</sup>是专门用于小样本计数训练测试的公 开数据集,也是目前小样本计数领域唯一的公开数据 集。该数据集包含 147 个类别的物体,拥有 6 135 张 图像。每张图像有 3 个样本来描述目标对象,并用矩 形框标注出来。需要注意的是,为了验证模型效果与 类别无关,训练集、验证集和测试集的目标物体类别 之间没有交集。训练集包含 89 个类,验证集和测试集 各包含互不关联的 29 个类。数据集中每张图像被计数 的目标物体数量差异较大,从 7 到 3 731 不等,平均 数为 56。

为了验证模型的通用性,虽然本模型是不特定类 别物体计数模型,但也可以在特定类别计数数据集中 进行训练测试。CARPK<sup>[2]</sup>是车辆计数的数据集,提供 了多个不同场景下的车辆图像和标注信息,用于训练 和测试车辆检测与计数算法的性能。CARPK 数据集包 含大量的车辆图像,涵盖了不同的环境和场景,如室 外停车场、街道和人工停车场等。这使得算法可以在 不同的情况下进行训练和测试,从而提高泛化能力。 一些图像可能存在遮挡、不同角度的车辆、光照变 化等因素,增加了算法处理的难度。通过使用这些 具有挑战性的图像,可以评估算法在复杂场景下的 鲁棒性和准确性。本文的模型在 CARPK 数据集中选 取的输入与之前特定类别计数模型不同。小样本计数 模型不仅需要输入查询图像,还需要输入样本图像, 因此对该数据集随机抽取3个样本输入模型进行训练 测试。

#### 3.2 指标

深度学习指标是用于评估和衡量深度学习模型性能的度量标准。这些指标可以帮助了解模型在训练和测试过程中的表现,并比较不同模型之间的性能差异。本文选择平均绝对误差(Mean Absolute Error, MAE)和均方根误差(Root Mean Square Error, RMSE)来衡

量模型的性能。

MAE = 
$$\frac{1}{N_i} \sum_{i=1}^{N_i} \left| C_i - C_i^{\text{gt}} \right|$$
 (6)

RMSE = 
$$\sqrt{\frac{1}{N_i} \sum_{i=1}^{N_i} (C_i - C_i^{\text{gt}})^2}$$
 (7)

式中, *N<sub>i</sub>* 是查询图像的数量; *C<sub>i</sub>* 和 *C<sup>gt</sup>* 分别是第*i* 个 查询图像的预测计数与真实计数。

#### 3.3 实验结果

本文的模型在 FSC147 数据集上部分图片的可视 化结果如图 2 所示。可以看到,针对不同类别的物体, 该模型在密度图估计上有良好的效果。

该模型在不同场景下的优势与劣势如图 3 所示。 在面对由于物体远近导致的样本与目标物体大小不一 致、模糊等问题时,该模型有较好的效果,如图 3 (a) 所示。该模型能处理由于视点远近所导致的问题,远 处的目标物体也能被成功检测到并计数。但在面对眼 镜这类语义较为复杂的物体时,该模型可能会把一副 眼镜识别为 2 副,如图 3 (b)所示。这主要是因为一 副眼镜有 2 个镜片,在特征增强时,过度强调了镜片 的局部属性,而忽视了其整体结构,因此导致计数值 不准确。



图 2 密度图预测结果 Fig. 2 Prediction results of density map



(a) 优势

(b) 劣势

#### 图 3 该模型在不同场景下的优势与劣势 Fig. 3 Advantages and disadvantages of the model in different scenarios

各模型在 FSC-147、CARPK 数据集中的比较分别 如表 1 和表 2 所示,其中 Val 表示验证集,Test 表示 测试集。

表1 各模型在 FSC-147 数据集中的比较 Tab.1 Comparison of methods on the FSC-147 dataset

算法	Val MAE	Val RMSE	Test MAE	Test RMSE
GMN <sup>[13]</sup>	29.66	89.81	26.52	124.57
FamNet <sup>[14]</sup>	24.32	70.94	22.56	101.54
FamNet+ <sup>[14]</sup>	23.75	69.07	22.08	99.54
本文所提模型	15.93	57.46	15.97	95.71

表 2 各模型在 CARPK 数据集中的比较 Tab. 2 Comparison of methods on the CARPK dataset

算法	MAE	RMSE
YOLO <sup>[5]</sup>	48.89	57.55
F-RCNN <sup>[17]</sup>	47.45	57.39
RetinaNet <sup>[6]</sup>	16.62	22.30
GMN <sup>[13]</sup>	7.48	9.90
FamNet <sup>[14]</sup>	18.19	33.66
本文所提模型	5.71	7.72

## 3.4 消融实验

本节研究了多尺度和特征增强模块对网络性能的 影响。具体而言,去除多尺度(Feature Enhancement only, FE-only)时,只在Restnet 50的第三层进行特 征增强计数。去除特征增强模块(U-Net only)时,计 算相似图,将样本特征、相似图和查询图像特征拼接 后,再送入回归器中计数。

各模块在 FSC-147 数据集上的消融实验结果如表 3 所示。实验证明,引入多尺度和特征增强能有效提高网络性能。

表 3 各模块在 FSC-147 数据集上的消融实验结果 Tab. 3 Experimental results of module ablation on the FSC-147 dataset

算法	Val MAE	Val RMSE	Test MAE	Test RMSE
U-Net only	21.69	63.55	21.53	99.53
FE only	16.23	59.71	17.19	96.32
U-Net 和 FE	15.93	57.46	15.97	95.71

#### 4 结论

本文构建了以 U-Net 结构为基础的多尺度网络, 在 3 个尺度上分别进行样本特征增强。以样本和查询 图像的内积相似度作为权重,将样本特征卷积相似图, 得到与查询特征图大小一致的样本增强特征图。将其 与原查询图像相加后进行上采样,送入上一级特征图 中。在各个尺度上都利用回归器进行计数,得到各级 密度图,相加求和后得到密度图。对密度图求和即可 得到计数总值。本文所提模型在 FSC-147 和 CARPK 数据集中进行训练和测试,实验结果表明,该模型的 性能优于大多数其他方法,能较好地处理由于视点远 近所导致的问题,但在处理眼镜这类既强调局部特征 又关注整体结构的物体时,效果欠佳。如果过分关注 整体结构,可能会导致模型难以处理由于遮挡导致的 非刚性形变问题。如何平衡两者之间的关系,则是小 样本计数的一个重难点。

#### 参考文献

 LIU L, CHEN J, WU H, et al. Cross-modal collaborative representation learning and a large-scale rgbt benchmark for crowd counting[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and

- [2] HSEH M R, LIN Y L, HSU W H. Drone-based object counting by spatially regularized regional proposal network[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2017: 4145-4153.
- [3] 杜妍辰,孙洁,汪晓铭,等. 一种基于表面肌电信号映射人体下肢运动意图的方法[J]. 生物医学工程学进展,2023,44(2):158-162.
  DU Yanchen, SUN Jie, WANG Xiaoming, *et al.* A method of human lower limb motion intention mapping based on surface electromyography[J].
  Progress in Biomedical Engineering, 2023, 44(2): 158-162.
- [4] 宗西增,蔡蕊蕊,田若婷,等.基于多层感知机的 DNA 甲基化年 龄预测模型[J]. 生物医学工程学进展, 2023, 44(1): 34-41. ZONG Xizeng, CAI Ruirui, TIAN Ruoting, *et al.* DNA methylation age prediction model based on multilayer perceptron[J]. Progress in Biomedical Engineering, 2023, 44(1): 34-41.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 779-788.
- [6] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2017: 2980-2988.
- [7] DUAN K, BAI S, XIE L, et al. CenteNnet: keypoint triplets for object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2019: 6569-6578.
- [8] LU H, DAI Y, SHEN C, et al. Indices matter: learning to index for deep image matting[C]//Proceedings of the IEEE/CVF International

Conference on Computer Vision. IEEE/CVF, 2019: 3266-3275.

- [9] LU H, DAI Y, SHEN C, et al. Index networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(1): 242-255.
- [10] MA Z, WEI X, HONG X, et al. Bayesian loss for crowd count estimation with point supervision[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE/CVF, 2019: 6142-6151.
- [11] ZENG L, XU X, CAI B, et al. Multi-scale convolutional neural networks for crowd counting[C]//2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017: 465-469.
- [12] WANG Y, HU S, WANG G, et al. Multi-scale dilated convolution of convolutional neural network for crowd counting[J]. Multimedia Tools and Applications, 2020, 79(1): 1057-1073.
- [13] LU E, XIE W, ZISSERMAN A. Class-agnostic counting[C]//Asian Conference on Computer Vision. Springer, Cham, 2018: 669-684.
- [14] RANJAN V, SHARMA U, NGUYEN T, et al. Learning to count everything[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE/CVF, 2021: 3394-3403.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [16] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015: 234-241.
- [17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91-99.