

# 基于改进 SSD 的少样本目标检测

毕忠勤<sup>1</sup>, 单美静<sup>2</sup>, 刘志斌<sup>1</sup>, 徐富强<sup>1</sup>

(1. 上海电力大学 计算机科学与技术学院, 上海 200090;

2. 华东政法大学 信息科学与技术系, 上海 201620)

**摘要:**目标检测作为深度学习的热点问题之一,在自动驾驶、行人识别、智能医疗、机器人视觉等多个领域有着广泛的应用前景。但现有的大部分目标检测模型都依赖于大规模的标注数据集来训练模型以保证目标检测的准确率,而在许多实际的应用场景中,大量数据的标注不仅耗费人力物力,而且需要大量专业人士的参与,在一定程度上限制了目标检测模型的实际应用。针对少样本目标检测的特殊要求,基于 SSD 网络提出了一种改进的少样本目标检测模型,提高了目标检测应用的适用性。首先,在 SSD (Single Shot multiBox Detector) 网络的基础上,用 ResNet-50 代替 VGG 作为特征网络,从而提升模型的特征提取能力。其次,通过引入残差单元避免了网络退化问题。最后,为了充分融合各层之间的语义信息和位置信息,用 FPN (Feature Pyramid Networks) 替换了原模型中间的两个特征层。基于改进 SSD 网络的目标检测模型在少样本数据集的检测结果中, mAP 值达到了 79.8%, 比原始模型提高了 2.6 个百分点。

**关键词:**目标检测;机器学习;少样本;FPN;SSD 网络模型

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2023)11-0035-06

doi:10.3969/j.issn.1673-629X.2023.11.006

## Few-shot Object Detection Based on Improved SSD

BI Zhong-qin<sup>1</sup>, SHAN Mei-jing<sup>2</sup>, LIU Zhi-bin<sup>1</sup>, XU Fu-qiang<sup>1</sup>

(1. School of Computer Science and Technology, Shanghai University of Electric Power,  
Shanghai 200090, China;

2. Department of Information Science and Technology, East China University of  
Political Science and Law, Shanghai 201620, China)

**Abstract:** As one of the hot topics of deep learning, object detection has a wide application prospect in many fields, such as automatic driving, pedestrian recognition, intelligent medical treatment, and robot vision and so on. However, most of the existing object detection models rely on large-scale annotation data sets for model training to ensure the accuracy of target detection. In many practical application scenarios, the annotation of a large number of data not only consumes human and material resources, but also requires the participation of a large number of professionals, which limits the practical application of the object detection model to a certain extent. Aiming at the special requirements of few-shot object detection, we propose an improved few-shot object detection model based on SSD (Single Shot multiBox Detector) network, which improves the applicability of object detection applications. Firstly, the ResNet-50 is used instead of VGG as the feature network on the basis of SSD network to improve the feature extraction capability of the model. Secondly, the problem of network degradation is avoided by introducing residual element. Finally, in order to fully integrate the semantic information and location information between the layers, FPN is used to replace the two feature layers in the middle of the original model. In the detection results of the target detection model based on the improved SSD network in a few-shot data set, the mAP value of the improved model reached 79.8%, which was 2.6 percentage points higher than that of the original model.

**Key words:** object detection; machine learning; few-shot; FPN; SSD

## 0 引言

目标检测任务是计算机视觉领域最重要的任务之一,其研究主要用于人脸检测、安全系统、行人检测、自动驾驶、车辆检测以及医疗领域等<sup>[1]</sup>。在目前已知的

研究领域中,要训练一个好的模型往往需要大量的数据,而人工智能令人印象深刻的特征是能够仅通过一个或几个样本就快速建立对新颖概念的认知的能力。这意味着当前的目标检测技术还有很大的发展空间。

收稿日期:2022-12-07

修回日期:2023-04-11

基金项目:上海市地方能力建设项目(23010501500)

作者简介:毕忠勤(1977-),男,博士,教授,研究方向为人工智能、计算机视觉;通讯作者:刘志斌(1989-),男,硕士,研究方向为计算机视觉。

为了提高目标检测应用的适用性,少样本目标检测任务成了当前研究的重点<sup>[2]</sup>。涉及到少样本目标检测的研究起步较早,早期的检测方法主要有基于模式匹配的方法,即使用小部分的样本作为模板,使用该样本在测试样本中寻找目标,后来也出现过形状匹配算法,即通过制定的标准来检测目标之间的相似性,然而传统的目标检测算法往往运算量大并且速度缓慢<sup>[3]</sup>。传统方法的特点是候选区域几何特征的选择,直线特征识别是早期图像目标的基本特征之一,其检测算法主要包括:标准霍夫变换(SHT)<sup>[4]</sup>、Burns<sup>[5]</sup>方法和 LSD<sup>[6]</sup>方法等。由于霍夫变换对随机噪声和特征部分的遮挡具有较高鲁棒性,该方法得到了广泛的运用。而多边形特征识别是通过候选区域的各边界端点的图像二维坐标来进行描述,通过对各相邻特征点连线即可恢复完整的多边形信息。该方法的问题是,由于空间环境问题使得边界特征不完整,从而造成边界特征描述上的困难。传统的目标检测算法取得了很大的进展,但是仍然存在一些问题,例如,基于滑动窗口的区域选择策略没有针对性,时间复杂度高,窗口冗余,手工设计的特征对于环境多样性的变化并没有很好的鲁棒性。

随着深度学习的发展,基于深度学习的少样本目标检测逐渐成为了研究热点。根据监督机制的不同,基于深度学习的少样本目标检测算法主要分为有监督少样本目标检测、半监督少样本目标检测以及弱监督少样本目标检测。有监督少样本目标检测主要采用元学习方法、迁移学习方法或者数据增强方法。半监督

少样本目标检测则主要基于自我训练或自我监督两种解决方案。弱监督少样本目标检测则主要包括基于主动对抗的方法、多实例学习方法以及基于度量学习的方法。

该文对 SSD<sup>[7]</sup>网络运用于少样本目标检测中存在的问题进行分析,优化 SSD 网络,优化后的模型不仅可以提高主干网络的特征提取能力,还可以有效提升对少样本数据集检测的准确性。首先将 SSD 网络中的主干网络替换成层数更深的 ResNet-50,高层次的网络模型有效地提高了特征提取能力,并且利用残差单元避免了网络退化问题。其次,用 FPN 模块替换了原模型中的两个特征层,充分地融合了语义信息和位置信息。

## 1 SSD 网络

SSD 网络是传统的 one stage 目标检测算法,SSD 的特点是生成的候选框都是相同尺寸,也是评分机制对候选框内的目标物体进行预测,最后使用非极大值抑制产生最终的检测结果。

SSD 网络结构如图 1 所示。首先数据从输入层输入到模型中,接着由特征提取网络提取数据特征。SSD 的特征提取网络是由原始的 VGG16 网络和另外的五个卷积层构成,增强了对复杂特征的提取能力。从图中可以直观看出这些层的尺寸不一样,是为了实现在不同的尺度上对特征图进行预测。

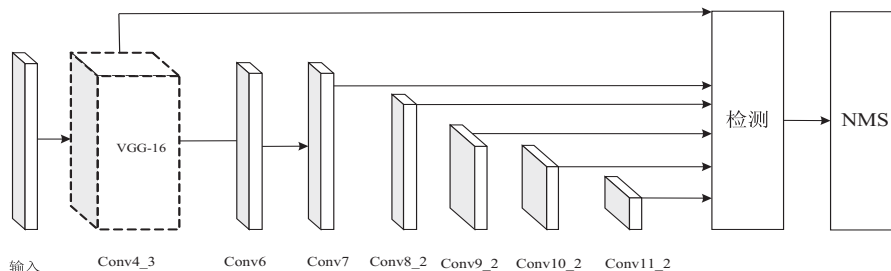


图 1 SSD 网络模型

SSD 算法中,提出了一个新的概念,即先验框,在特征提取网络生成特征图之后,以每个特征图为中心都会产生一系列不同尺度的先验框,这些先验框与原图一一对应,通过损失函数检测先验框的位置与真实目标框的位置的重叠度,从而分析出这个先验框内目标物体的类别,而且通过检测,也会对先验框的形状进行调整,使其大小和尺度与真实的目标框更贴合。

先验框包含两个参数:尺度  $S$  和宽高比  $a$ 。假设特征提取网络总共产生  $m$  个特征图,那么每个特征图先验框尺度的计算方式如公式(1)所示:

$$S_k = S_{\min} + \frac{S_{\max} - S_{\min}}{m - 1}(k - 1), k \in [1, m] \quad (1)$$

SSD 算法对不同特征层生成特征图的尺度制定了如下规则:先验框的尺寸随着网络层数不断增加而增加,而特征图尺寸逐渐减小,不同特征层生成特征图的详细尺寸及先验框的大小如表 1 所示。

表 1 每个特征图上的初始框尺度

Index	Feature map	Feature map size	Default box scale	True size
1	Conv6	38	0.2	60
2	Conv7	19	0.34	102
3	Conv8_2	10	0.48	144
4	Conv9_2	5	0.62	186
5	Conv10_2	3	0.76	228
6	Conv11_2	1	0.9	270

对先验框的长宽比采用公式(2)的设置方式:

$$a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\} \quad (2)$$

每个特征图上的初始框宽、高可以用公式(3)得出:

$$w_k^a = s_k \sqrt{a_r}, \quad h_k^a = s_k / \sqrt{a_r} \quad (3)$$

对于宽高比为 1 的初始框,又多增加了一个正方形的初始框,定义如公式(4)所示。

$$s_k = \sqrt{s_k s_{k+1}} \quad (4)$$

SSD 所用到的特征图及其大小如表 2 所示。

表 2 SSD 网络所用到的特征图及其大小

feature map	Size
Conv4_3	38 * 38
Conv7	19 * 19
Conv8_2	10 * 10
Conv9_2	5 * 5
Conv10_2	3 * 3
Conv11_2	1 * 1

为了得到更好的检测结果,SSD 引入 ground truth 的概念,即那些预测正确的样本。算法的关键是要搞清楚用哪个候选框进行预测,SSD 的操作是用先验框和训练样本中的 ground truth 进行比对,如果匹配成功,那与之对应的候选框就负责预测它。所以先验框和 ground truth 的匹配规则是算法是否有效的关键因素。SSD 的先验框与 ground truth 的匹配规则可以分成以下两个步骤:

(1)对于图片中的众多候选目标,需要确定每个候选框对应的 ground truth,计算不同先验框和 ground truth 的 IOU,找出其中的最大值。这样就确定了一组匹配关系。因为 IOU 总有最大值,所以这种方式可以

保证每个 ground truth 都不会被遗漏。如果某个先验框和 ground truth 成功匹配,称之为正样本。反之,该先验框对应的候选框定位失败,只能是负样本。

(2)在第一轮筛选中剩余的先验框,不能全部舍弃,通过设定一个阈值,当某个 ground truth 与先验框的 IOU 大于这个阈值,那么也能达成匹配。但是这可能导致某个 ground truth 与多个先验框匹配成功,这是可以的,但是一个先验框只能匹配一个 ground truth。

## 2 改进 SSD 模型

SSD 算法的训练已经非常成熟,因为它不需要生成候选框并且简化了特征提取的过程。通过实验表明,SSD 在 VOC2007 数据集上的检测精度达到了 74.3%,比目前最先进的 Faster R-CNN 算法效果更好。与同类型的 one stage 算法相比,检测速度也更快。然而,SSD 仍然存在可以优化的地方:首先,该模型提出的先验框概念虽然有助于训练,但是先验框的初始尺度和长宽比需要手工设置,不能通过模型训练直接获得。而且 SSD 网络中每一层特征图使用的先验框的尺寸都不同,增加了训练的难度。其次,与 Faster R-CNN 算法相比,SSD 对小样本的检测效果还存在明显的差距。因为 SSD 的底层特征图结构简单,无法对小目标的特征进行充分的提取。针对这些问题,该文对 SSD 算法进行优化,改进后的网络模型如图 2 所示,而对 SSD 算法提出的改进主要有:

(1)将 SSD 的特征网络 VGG 替换成网络层数更深的 ResNet-50。增强对复杂特征的提取,并且利用残差单元,避免了网络层数加深带来的退化问题。

(2)将 SSD 结构中 Conv8\_2 和 Conv9\_2 两个特征层替换成 FPN,更充分地融合不同层的语义信息,提升检测效果。

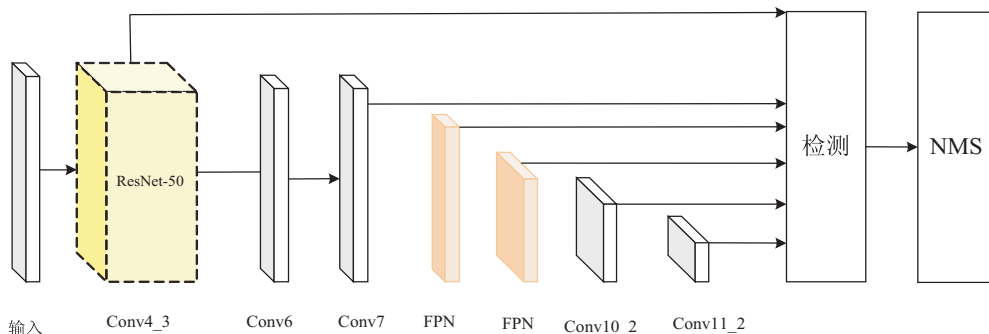


图 2 改进的 SSD 模型结构示意图

下面对图 2 中的 ResNet-50、FPN 模块以及损失函数进行介绍。

### 2.1 ResNet 模型

ResNet<sup>[8]</sup>(残差网络)是由何凯明等提出的,解决了深层次网络模型退化的问题。ResNet 模型是在

VGG19 模型的基础上进行了修改,该算法提出了残差单元这个概念。残差单元网络结构如图 3 所示。假设残差网络的输入为  $X$ ,将学习到的特征记为  $H(X)$ ,使残差  $F(X) = H(X) - X$ ,那么就会得到原始的特征为  $F(X) + X$ 。当残差为 0 时,即使使用多个网络堆积到

一起进行训练,也能保证网络性能不会下降,实际上残差不可能为 0,那么深层次的网络结构就可以继续提取新的特征,从而拥有更好的性能。ResNet 还有一个重要的训练原则:当特征图大小降低一半时,它的数量增加一倍,保持了网络层的复杂度。

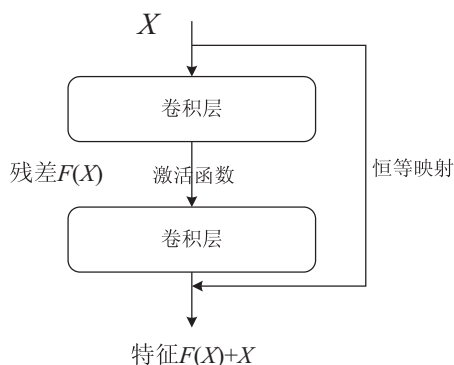


图 3 残差单元网络示意图

## 2.2 FPN 模块

传统的 Faster R-CNN<sup>[9]</sup> 网络集成了 R-CNN 系列算法的优点,逐步解决了 R-CNN 系列算法存在的问题,但是该算法在小样本检测任务中表现不是很好,原因在于无论是 RPN<sup>[10]</sup> 网络还是 Fast R-CNN<sup>[11]</sup> 网络,都是基于单个高层特征进行任务提取,那么小目标经过多次卷积池化,特征层中的小目标特征已经所剩无几。解决这种问题的常用方法是采用图像金字塔的方式在训练或测试阶段对图片进行多尺度变化增强,但是这样带来了极大的计算量。FPN 主要解决的是物体检测中的多尺度问题,通过简单的网络连接改变,在基本不增加原有模型计算量的情况下,大幅度提升了小物体检测的性能。FPN<sup>[12]</sup> 的提出是为了结合高层特征和低层特征,借鉴了 ResNet 的跳接结构,使得待检测的特征图中具有丰富的语义信息和位置信息。并且,该模型利用多尺度特征图检测,较好地处理了目标检测中的多尺度变化问题。

## 2.3 损失函数

损失函数一般用来评价模型的预测值和真实值的误差,对网络学习速度的快慢及最终模型预测效果的好坏起着关键的作用。IOU<sup>[13]</sup> 是目标检测中常用的指标,用来反映预测框和目标框之间的检测效果,其定义为:

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

SSD 中用 CIOU<sup>[14]</sup> 代替原始的 BBOX 的回归损失函数。对于目标检测来讲,在获得真实框之后需要先对真实框进行编码,将其转化为预测框的形式,然后将真实框与网络的预测结果进行对比,对网络结构进行优化,使框的预测更准确。CIOU 是 IOU 的一个改进

版,公式如下:

$$CIOU = IOU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha\nu \quad (6)$$

式中,  $b$ 、 $b^{gt}$  分别表示预测框和真实框的中心点;  $\rho^2$  表示两个中心点的欧氏距离;  $c$  表示能够同时包含预测框和真实框的最小闭包区域的对角线距离。其中:

$$\alpha = \frac{\nu}{1 - IOU + \nu} \quad (7)$$

$$\nu = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

$\alpha\nu$  与真实框和预测框的框高有关系,二者越相近,  $\alpha\nu$  的值越低, CIOU 指的是真实框和预测框的偏离程度,则 LOSS 函数为:

$$LOSS_{CIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha\nu$$

## 3 实验及分析

### 3.1 公共基准数据集介绍

该文采用 FSOD 数据集作为测试数据集。FSOD 数据集是专为少样本学习和评估新颖类别模型的通用性而设计的,该模型包含 1 000 个类别,分别用于训练和测试集的 80/20 分割,总共约 66 000 张图像和 182 000 个候选框。该数据集包含 83 种父级语义,例如哺乳动物、衣服、武器等,并进一步划分为 1 000 个叶子类别。数据集包含对象大小和纵横比差异很大的对象,由 26.5% 的图像组成,其中测试集中的对象不少于 3 个。测试集包含大量未包含在标签系统中的类别的候选框<sup>[15]</sup>。

### 3.2 实验环境及步骤

该文进行的实验是在 Ubuntu16.04 操作系统下完成的。实验所采用的硬件配置为:中央处理器(CPU): Intel Core i9-10900k; 图像处理器(GPU): NVIDIA GTX 3080 独立显卡, 10G 显存。使用的编程语言为 Python3.6, 机器学习开发框架为 Pytorch1.7。

实验步骤如下:

(1) 搭建文中网络。

(2) 使用 ResNet-50 在 Mini-ImageNet 分类数据集进行预训练,得到预训练后的 ResNet-50 模型。

(3) 进行迁移学习,去除 ResNet-50 模型的全连接层,加载其权重作为文中主干网络的初始参数,对类别预测分支、位置预测分支使用 kaiming 分布进行参数初始化,对其余部分参数随机初始化。

(4) 加载 Mini-ImageNet 训练集图片,将图片大小设置为 224×224,使用随机裁剪、缩放、翻转等方法进行数据增广,输入网络进行训练。

(5) 使用随机梯度下降法作为优化器,初始学习率设置为 0.001, batch\_size 大小设置为 8,最大迭代次

数设置为 200。优化器使用 Adam,其参数为 step\_size = 1, gamma=0.95。

(6)保存训练模型。

(7)加载训练完毕的模型,输入 Mini-ImageNet 测试集图片,输出测试结果。

### 3.3 评价指标

在目标检测领域,往往使用平均精度均值 (mean Average Precision, mAP)、平均精度 (Average Precision, AP)、精确率 (Precision) 以及召回率 (Recall) 来评估模型的检测效果;使用每秒传输帧数 (Frame Per Second, FPS)、每秒 10 亿的浮点运算次数 (Giga Floating-point Operations Per Second, GigaFLOPS) 来评估模型的检测性能。该文采用 FPS 和 mAP 两个指标衡量改进算法的效率。公式 (9) 和公式 (10) 为 Precision 和 Recall 的计算公式。

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (9)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (10)$$

式中,TP 代表模型认为是正样本但实际上是正样本的输入,FP 代表模型认为是正样本但实际上不是正样本的输入,FN 代表模型认为是负样本但实际上不是负样本的输入。

由 Precision 和 Recall 作为横纵坐标可以得到一条 PR 曲线,PR 曲线下的面积则被称为 AP 指标,对精确率和召回率的结果进行了综合评估。定义如下:

$$\text{AP}_i = \sum_{k=1}^N p(k) \Delta r(k) \quad (11)$$

式中,  $p(k)$  为召回率变化点  $k$  对应的精确率;  $\Delta r(k)$  为变化点  $k$  对应的召回率的变化量;  $N$  为召回率变化点的数量;不同类别 AP 不同,  $i$  为类别的索引值。

mAP 指标则为所有的类的 AP 值进行求平均。定义如下:

$$\text{mAP} = \frac{1}{m} \sum_{i=1}^m \text{AP}_i \quad (12)$$

式中,  $m$  为类别数量。

### 3.4 图片数据预处理

(1)灰度世界算法是基于灰度世界假设的,灰度世界算法在物理上假设自然界物体光的平均反射值通常是一个固定值,它近似于灰色。采用色彩平衡方法将该假设应用于训练集中的图像,可以从图像中消除环境光的影响,从而达到增强图像的效果。

(2)通过对训练样本旋转来实现数据扩增。旋转包括水平旋转、垂直旋转、镜像旋转,旋转的角度范围为  $+30^\circ \sim -30^\circ$ ,并对新增样本进行筛选,去除劣质样本。

(3)数据集的制作。在经过对图像数据的预处理之后,得到的数据如表 3 所示,分为三类。第一类是原

始图像训练集,第二类是经过图像增强后的训练集,第三类是经过样本扩增后的训练集。

表 3 训练集数量

训练集	方法	图片数量
1	原始数据	48 000
2	图像增强	48 000
3	样本扩增	143 092

### 3.5 实验结果和对比分析

首先,为了验证提出的改进算法的有效性,对改进的 SSD 网络模型与 SSD 网络模型进行实验对比。

如表 4 所示,改进后的模型采用 ResNet-50 作为主干网络,经过实验,算法在 Mini-ImageNet 数据集上的 mAP 有了明显的提升。这是因为 ResNet-50 的网络层数更深,加强了对复杂特征的提取,提高了检测效果。

表 4 改进的算法和原始 SSD 算法的平均精度对比

算法	主干网络	mAP/%
SSD <sup>[7]</sup>	VGG	77.2
改进后的算法	ResNet-50	79.8

表 5 展示了所提算法与 SSD 系列改进算法 DSOD、RSSD、DSSD、FSSD 的详细对比。从表中可以看出 SSD300 的 FPS 最高,达到了 46。所提改进算法在检测速度上对比 SSD300 有所下降,但是在输入图像分辨率、准确率这两个指标上达到了最优,改进后的算法在速度没有明显降低的情况下 mAP 达到了 79.8%。

表 5 所提算法与 SSD 系列改进算法的对比

算法	主干网络	图片大小	速度 /FPS	mAP /%
SSD300 <sup>[10]</sup>	VGG	300 * 300	46	77.2
DSOD300 <sup>[16]</sup>	VGG	300 * 300	17.4	77.7
RSSD300 <sup>[17]</sup>	DS/64-192-48-1	300 * 300	35	78.5
DSSD321 <sup>[18]</sup>	ResNet-101	321 * 321	9.5	78.6
FSSD300 <sup>[19]</sup>	VGG	300 * 300	35.6	75.3
Ours	ResNet-50	224 * 224	38.3	79.8

## 4 结束语

该文提出了一种基于 SSD 算法的少样本目标检测算法,针对 SSD 算法在少样本目标检测中的问题,提出了改进思路。首先,用神经网络层数更深的 ResNet-50 替换原来的 VGG,ResNet-50 作为主干网络,可以对更加复杂的特征进行提取,并且利用残差单元避免了随着网络层数加深带来的退化问题;其次,用 FPN 替换 SSD 网络中间的两个特征层,充分融合不同

层次的特征信息。通过实验,改进的 SSD 算法在 Mini-ImageNet 数据集的检测结果中, mAP 值达到了 79.8%, 高于原始的 SSD 算法及其系列的改进算法, FPS 也达到了 38.3。实验结果表明提出的改进算法是有效的,并且具有良好的鲁棒性。另外,为了增强少样本检测的检测效果,提出了一种图像处理的方法,该方法会增加一定的计算量,如何设计出更理想的图像处理方案或者引入更适合研究的少样本数据集,这是后续研究值得考虑的问题。

#### 参考文献:

- [1] 谢 富,朱定局.深度学习目标检测方法综述[J].计算机系统应用,2022,31(2):1-12.
- [2] 董文轩,梁宏涛,刘国柱,等.深度卷积应用于目标检测方法综述[J].计算机科学与探索,2022,16(5):1025-1042.
- [3] HOSPEDALES T M, ANTONIOU A, MICAELLI P, et al. Meta-learning in neural networks: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9):5149-5169.
- [4] KIM G, JUNG H, LEE S, et al. Few-shot object detection via knowledge transfer[C]//2020 IEEE international conference on systems, man, and cybernetics (SMC). Toronto: IEEE, 2020:3564-3569.
- [5] 林 杰.基于深度学习的目标检测系统的研究与实现[D].北京:北京邮电大学,2021.
- [6] 周奇丰,凌莉萍.基于 OpenCV 视觉库的车道线图像识别[J].汽车实用技术,2021,46(17):18-21.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//Computer vision - ECCV 2016. Amsterdam: Springer, 2016:21-37.
- [8] SHIBATA N, TANITO M, MITSUHASHI Ket al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography[J]. Scientific Reports, 2018, 8(1):14665.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [10] FAN Q, ZHUO W, TANG C K, et al. Few-shot object detection with attention-RPN and multi-relation detector [C]//2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Seattle: IEEE, 2020:4012-4021.
- [11] GIRSHICK R. Fast R-CNN [C]//2015 IEEE international conference on computer vision. Santiago: IEEE, 2015:1440-1448.
- [12] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//30th IEEE conference on computer vision and pattern recognition (CVPR 2017). Honolulu: IEEE, 2017:936-944.
- [13] CHOWDHURY R R, BATHULA D R. Influential prototypical networks for few shot learning: a dermatological case study [C]//2022 IEEE international symposium on biomedical imaging (IEEE ISBI 2022). Kolkata: IEEE, 2022.
- [14] MUNKHDALAI T, YU H. Meta networks [J]. arXiv:1703.00837, 2017.
- [15] LAURA B, MICHAEL B. Are convolutional neural networks trained on ImageNet images wearing rose-colored glasses?: a quantitative comparison of ImageNet, computed tomographic, magnetic resonance, chest X-Ray, and point-of-care ultrasound images for quality [J]. Journal of Ultrasound in Medicine, 2020, 40(2):377-383.
- [16] SHEN Z. DSOD: learning deeply supervised object detectors from scratch [C]//Computer vision and pattern recognition. Venice: IEEE, 2017:1937-1945.
- [17] ZHOU S, QIU J. RSSD: object detection via attention regions in SSD detector [C]//2019 2nd international conference on safety produce informatization, IICSPI 2019. Chongqing: [s. n.], 2019:266-269.
- [18] FU Chengyang, LIU Wei, RANGA A, et al. DSSD: deconvolutional single shot detector [J]. arXiv:1701.06659, 2017.
- [19] LI Z, ZHOU F. FSSD: feature fusion single shot multibox detector [J]. arXiv:1712.00960, 2017.