

潘兴甲^{1,2} 张旭龙^{1,2} 董未名¹ 姚寒星³ 徐常胜^{1,2}

小样本目标检测的研究现状

摘要

近年来,随着深度学习技术的进步与推广,目标检测领域得到快速发展.但目前基于深度学习的方法大多对大规模标注数据有着极高的需求,而现实场景中大量标注数据往往不可能.因此,基于少量标注样本的目标检测领域逐渐得到大家关注.本文系统地总结与分析是目前有关小样本目标检测的方法,指出了目前方法的缺陷,并提出了一些可能的发展方向.

关键词

小样本学习;深度学习;目标检测

中图分类号 TP181

文献标志码 A

收稿日期 2019-10-17

资助项目 国家重点研发计划(2018YFC0807500)

作者简介

潘兴甲,男,博士生,研究方向为机器学习与计算机视觉.xingjia.pan@nlpr.ia.ac.cn

董未名(通信作者),男,博士,研究员,研究领域包括计算机图形学、机器学习与计算机视觉.weiming.dong@ia.ac.cn

1 中国科学院自动化研究所,北京,100190

2 中国科学院大学,北京,100190

3 北京亮亮视野科技有限公司,北京,100014

0 引言

目标检测是计算机视觉的基础任务之一,主要任务是对图像中的目标进行分类和定位.早在 20 世纪 60 年代,目标检测技术已经被一些学者所关注,当时的研究方法主要基于几何学原理以及低层次的图像特征,诸如纹理、色彩等.近年来随着高性能并行计算技术的逐渐成熟以及大数据的积累,基于深度学习方法的目标检测技术开始涌现且迅速取代传统方法,有很多成熟的算法已经部署到实际场景中.目前较为成熟的方法大多借助于大规模的标注数据,导致应用场景单一、覆盖任务单一等.在绝大部分的应用场景中,收集满足要求的标注数据是一件十分耗费财力物力的工作,这极大限制了现有目标检测方法的落地应用与推广.基于此,应用较少标注数据的半监督方法与利用不完全匹配的标注数据的弱监督方法逐渐得到学者们的关注.更进一步,如何利用极少的标注数据学习具有一定泛化能力的模型是更加迫切需要研究的问题.

人类具有只看过一次或极少次某物体便可以永久记得的能力,即使是只有几岁的小孩子.基于深度学习的机器学习方法大多借鉴人类大脑的工作机制,尤其是卷积神经网络.面对同样的场景,小样本学习即研究如何利用极少标注样本进行有效的学习.近年来,小样本学习在图像分类领域取得了一些不错的成果,但小样本目标检测领域则鲜有研究.目标检测主要包含分类与回归两个任务,基于小样本的回归任务是极具挑战性的问题.本文首先系统地介绍小样本学习的方法,之后介绍小样本目标检测的研究进展,最后对已有的方法进行总结并提出了可能的研究方向.

1 小样本目标检测概念

1.1 小样本学习

深度学习目前广泛应用于图像识别、自然语言处理、视频分析等各个领域,并且取得了极大的成功,是典型的数据驱动类型技术.但实际应用场景中,由于标注数据的匮乏,简单而直接地应用深度学习技术往往不能取得很好的效果.小样本学习以此为出发点,利用少量标注数据,通过学习不同子任务中的共性,提高模型泛化性能,进而满足实际需求.小样本学习是元学习在监督学习领域的特例.目前方法大致分为 4 种:基于度量学习的、基于元学习的、基于数据增强的以及

基于多模态的方法.

1.2 小样本目标检测

一般的目标检测方法通常借助大量标注数据获得具有一定泛化性能的模型并在新数据中得以应用.例如被广泛应用的 OpenImages^[1]、ImageNet^[2]、MSCOCO^[3] 等数据集.而小样本目标检测面向的场景中通常包含大量的无标注数据与极少的标注数据.如图 1 所示,基于包含充足标注数据的基类,模型试图学习一定的泛化性能,并仅靠极少标注数据在新类上进行检测.融合传统的目标检测与小样本学习方法,通过巧妙的训练方法设计、模型结构设计与损失函数设计,引导模型在极少量标注数据中学习具有一定泛化性能的检测模型.

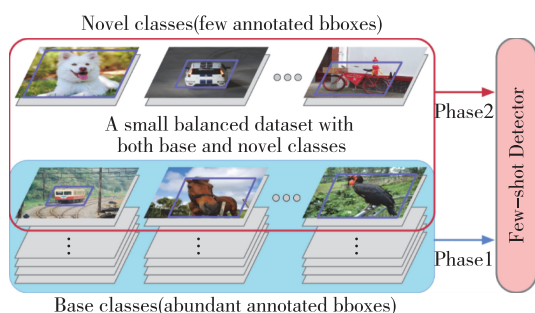


图 1 小样本目标检测示例^[37]

Fig. 1 An example of few-shot object detection^[37]

2 小样本目标检测方法

2.1 小样本学习介绍

小样本学习是元学习在监督领域的应用.在元训练阶段,通过组合不同的训练集构成不同的元任

务,模型得以学习独立于任务的泛化能力.在元测试阶段,面对全新的类别或者任务,不需要重新训练或者少量的快速训练即可完成识别任务.形式化地说,在元训练阶段,首先从训练集中随机抽取 C 个类别,每个类别包含 K 个样本,构成模型的支撑集,即为一个元任务.同时挑选少量图片作为测试集.目的为使得模型从 $C \times K$ 个数据中学会识别 C 个类别,此任务即为 C -way K -shot 问题.图 2 展示了 5-way 1-shot 的小样本学习实例.小样本学习中 K 值很小(通常 $K < 10$),当 $K = 1$ 时,称为单样本学习.小样本学习按照解决方案大致可分为 4 类:基于度量学习的方法、基于元学习的方法、基于数据增强的方法以及基于多模态的方法.图 3 所示展示了对应 4 种方法的典型网络结构.

1) 基于度量学习的方法

此类方法从改善框架角度入手.由于模型中包含数以万计的参数,直接按照普通的分类器基于交叉熵的方式去学习,最终结果不可避免地会遭遇过拟合问题.因此,基于度量学习的方法通过对样本间距离进行建模,借鉴传统非参数化的方法,使得同类目标距离相近、异类目标趋远.孪生网络^[4]包含权值共享的两个分支,在模型的顶部输出两个样本属于同一类别的概率.在测试阶段,模型输出测试样本与支撑集所有的样本属于同一类别的概率,取最高概率所属类别为最终结果.相比于孪生网络,匹配网络^[5]中支撑集和查询图像采用不同的特征提取器,分类结果由查询图像与支撑集所有样本预测值的加权和得到.文献[5]主要有两处改进:在建模过程中,提出了基于记忆与注意力机制的匹配网络,加快学习;在训练过程中,提出训练和测试需在同样条件下

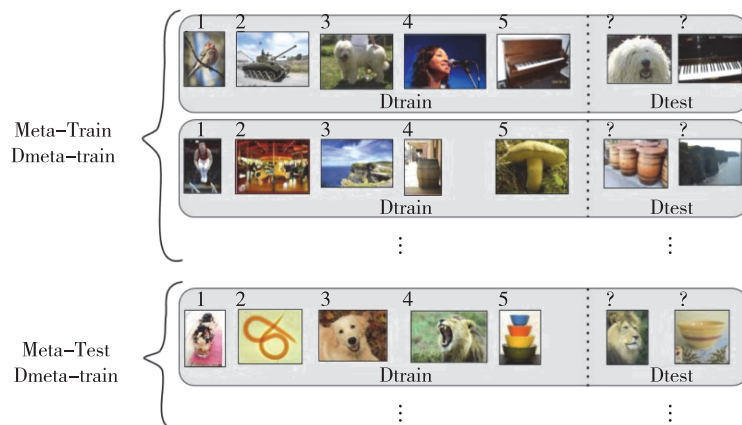


图 2 小样本学习示例^[14]

Fig. 2 An example of few-shot learning^[14]

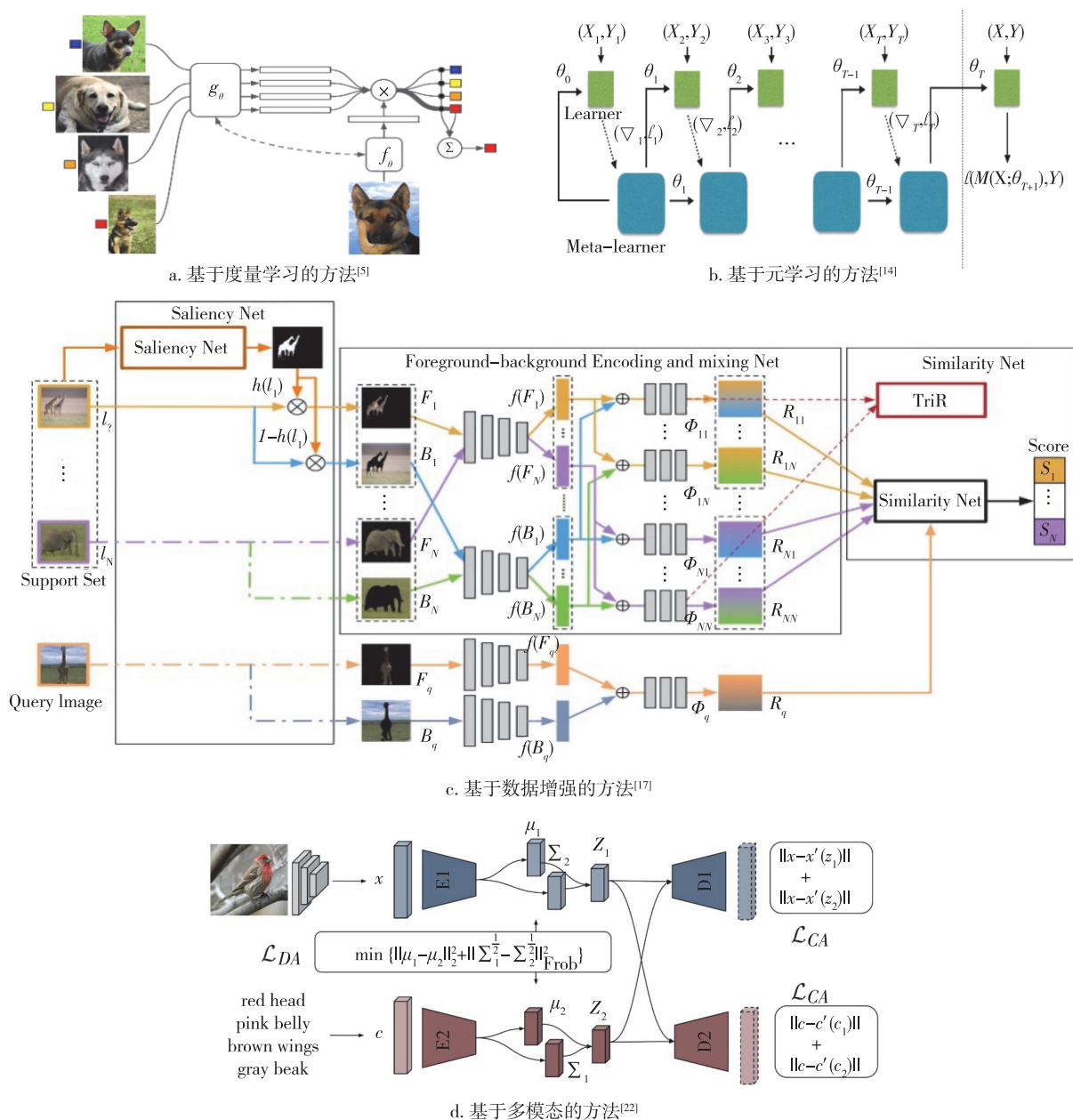


图3 4种不同的小样本学习方法

Fig. 3 Illumination of four few-shot learning methods^[5,14,17,22]

进行,即在训练阶段保证网络只能看到每一类的少量样本.另外,文献[5]采用LSTM(长短期记忆网络)使得所有图像得以交互,即完全上下文嵌入.为解决分类过程需比较查询图像与支撑集中所有样本效率较低的问题,原型网络^[6]提出每个类别都存在一个原型表达,某类的原型表达是支撑集在嵌入空间中的均值.因此,分类问题退化为比较嵌入空间中测试样本与各类原型的距离问题.更进一步,之前的方法大多采用简单且固定的距离度量方式,而Suang等^[7]认为距离度量本身也是一个值得学习的问题,

因此提出对距离度量进行建模的关系网络.网络主要包含2个模块,即编码模块与关系模块.通过关系模块在嵌入空间对距离度量进行学习,进而提升分类效果.最近的,Garcia等^[8]和Kim等^[9]通过图神经网络进行端到端地学习样本在特征空间的关系,利用图卷积操作更新样本特征及样本关系,取得了不错的效果.

2) 基于元学习的方法

Santoro等^[10]提出记忆增强网络解决小样本学习任务.该文借鉴神经图灵机可以快速编码及检索

新信息的思路,首先证明记忆增强网络可以迅速吸收新数据,并且利用少量数据即可做出精确的预测.另外,文献[10]提出一种新的专注于存储内容的外部存储索引的方法.具体的,文章基于 LSTM 等 RNN (递归神经网络)序列模型,将数据看作序列进行训练.测试时,输入新类别样本进行验证. Munkhdalai 等^[11]提出元网络,学习不同元任务中的元知识并通过快速参数化调整模型的推理偏置进而达到快速泛化的目的.模型包含两个主要的学习模块:元学习器旨在快速学习元任务之间的泛化信息,基学习器则是面向特定的任务.生成的快速权重同时集成到元学习器与基学习器中以调整学习器的归纳偏置. Jamal 等^[12]通过添加一个正则化项提出任务不可知的元学习器,使得模型在不同任务间的差距更小. Sun 等^[13]提出元迁移学习,主要包含两个部分,首先对预训练的模型进行调优,只学习每一层的放缩和偏置,之后进行困难任务挖掘.另外,一些研究者认为基于梯度优化的一系列优化方法并不能胜任小样本学习的任务. Ravi 等^[14]通过研究认为基于梯度的优化算法失败的原因在于:一是梯度优化算法大多无法在几步之内完成优化,特别是非凸问题,多种超参的选取降低了收敛速度;二是不同任务分别随机初始化会影响最终的收敛结果.该文提出学习模型参数的更新规则,而非参数.具体地,文中采用 LSTM 作为元学习器学习梯度下降的参数更新规则,目的是学会在新的类别上对分类器进行参数初始化与更新.与之不同, Finn 等^[15]提出模型无关的元学习方法,直接优化初始化,使得基于此初始化能采用小样本数据进行快速高效的收敛.该方法没有对模型形式做任何假设,且十分高效.

3) 基于数据增强的方法

在数据量不充足的情况下,进行数据增强是最朴素直接的方法. Alfassy 等^[16]通过在特征空间对输入特征对进行不同方式组合(求交、求并、做减)以产生更加丰富的特征样本.对抗生成网络是很好的数据增强工具,但同样需要大量的标注样本. Zhang 等^[17]利用预训练好的显著性图网络提取样本的前景与背景信息,之后分别经过独立的编码器,在特征空间通过融合网络进行重组,生成包含更多新生概念的样本. Chu 等^[18]考虑人在观察图像时随着视觉轨迹的变化,视线关注的区域同时在变化.文中通过强化学习模型预测轨迹,分别提取对应区域,再通过 RNN 序列模型聚合不同区域的特征得到分类结果,

其本质上是通过提取区域特征对图像样本进行了数据扩充. Chen 等^[19]首先提出不同程度的图像变形虽然会使得图像看上去不再真实,但对于分类问题仍然可以提供关键的语义信息.基于此,文献[9]提出包含图像变形子网络的元学习器,该学习器类似于 mixup^[20]的功能,可用于扩充原始少量样本.

4) 基于多模态的方法

在少量样本情况下引入其他模态信息进行融合同样是直观可行的. Schwartz 等^[21]提出对齐变分自编码器,模型包含两个分支,分别对图像与语义信息进行编码,通过多模态信息互助提高重建特征效果. Schonfeld 等^[22]同样引入类别的语言描述,借助多模态信息提高性能. Wang 等^[23]引入标签语义信息,通过嵌入损失迫使语义嵌入与视觉嵌入对齐,对于补充视觉特征的丰富性有很大帮助. Li 等^[24]引入标签语义信息,编码后以无监督的方式构建类别层次结构,这种层次分类的结构对于模型在未见过类的分类表现上更好.

2.2 小样本目标检测方法介绍

小样本目标检测是小样本学习方法在目标检测任务上的应用,旨在通过极少的标注样本学习具有较好检测性能的模型.小样本目标检测应用的数据通常包含具有丰富标注数据的源数据和标注数据极少的目标数据.小样本目标检测方法目前主要基于经典成熟的传统目标检测方法如 SSD^[25]、YOLO^[26]、Faster-RCNN 等^[27],借鉴小样本学习的算法框架,构建针对目标检测的小样本解决方案.现有的小样本目标检测方法主要可以分成 3 类:基于微调的方法、基于模型结构的方法以及基于度量学习的方法.图 4 展示了对应 3 种小样本目标检测方法的典型结构.

1) 基于微调的方法

实际场景中对于特定的目标检测任务,获取大规模的带有完整标注的数据集十分困难.在小样本学习领域,基于微调的迁移学习方法取得了较好的效果^[28-29]. Chen 等^[30]提出了 LSTD 的模型,该模型整合 SSD 与 Faster-RCNN,保留多尺度检测与两阶段检测框架.首先在包含大量标注数据的源域数据集上学习源域模型,并用该模型参数初始化目标域模型,之后在小规模数据上进行微调.简单的微调可能会忽略源域和目标域的重要目标知识,降低可传递性,为此文章中提出背景抑制正则化与知识迁移正则化,使 LSTD 在迁移学习过程中能够聚焦于前景目

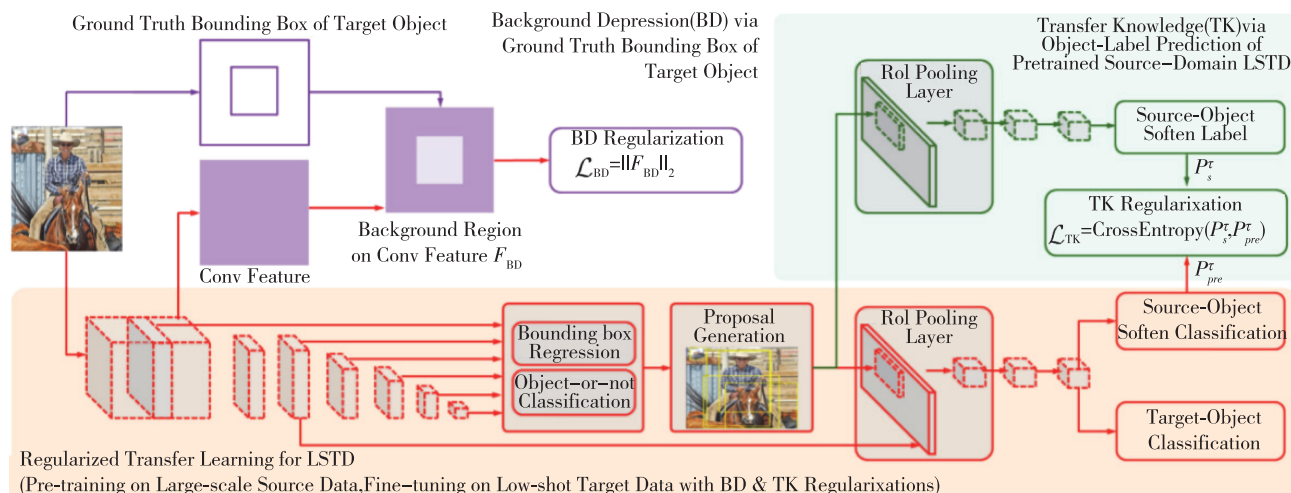
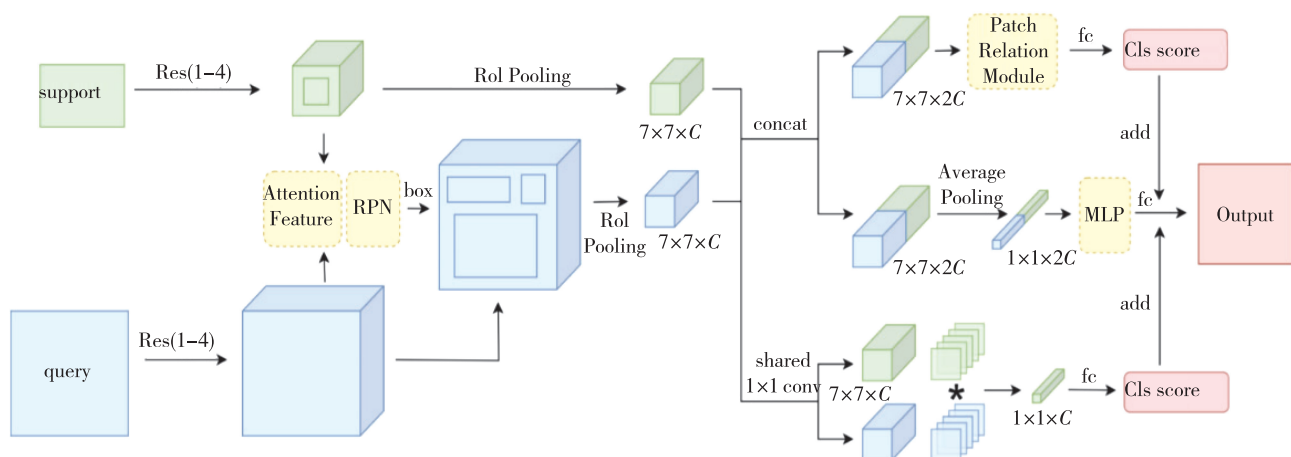
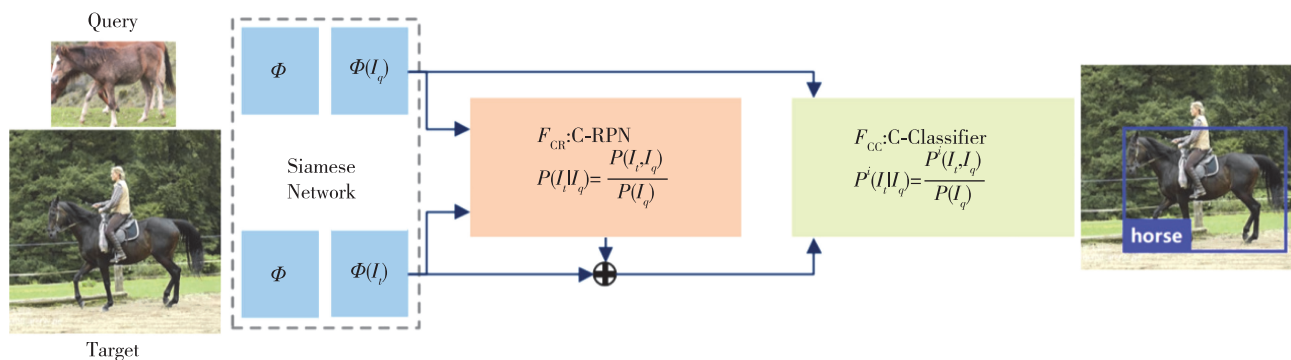
a. 基于微调的方法^[30]b. 基于模型的方法^[32]c. 基于度量学习的方法^[35]

图4 3种小样本目标检测方法示例

Fig. 4 Illumination of three few-shot object detection methods^[30, 32, 35]

标,减少背景干扰.Singh等^[31]同样采用微调的方法,借助LSTD模型通过在目标域微调预训练的模型得到检测结果.文献[31]首先通过目标检测网络对前景区域进行提取,之后借助纹理特征等训练更为精准的分类器.同样,他们还采用文献[30]提出的正则化技术,提高迁移学习过程中的泛化性能.

2) 基于模型的方法

一些研究者试图构建新的目标检测模型来解决小样本目标检测问题.Fan等^[32]研究发现检测未见过的类时,一般方法可能会完全错过目标或在背景上产生大量误检,认为可能是由于候选框分类置信度偏低所致.文章在区域候选框网络中引入注意力

机制,融合查询图像与支撑集图像的特征,向区域候选框网络中引入支撑集数据信息,以指导网络生成更多与查询图像相关的候选框并抑制无关的候选框.得到候选框后,在网络顶层检测器部分,模型中包含3条支路,每条支路均含有注意力模块.3条支路从局部、全局、互相关3个方面学习分别来自查询图像与支撑集的候选框的关系.最终的检测结果由3条支路共同作用得到.Dong等^[33]提出通过模型互交替优化的思路解决小样本目标检测的方法.该方法整合自主学习和多模型学习,提出多模型自主学习检测器.首先基于少量简单样本训练得到模型,再由现有模型得到可信赖并且更难的本,再进一步提升模型性能,如此循环往复.具体地,他们采用R-FCN^[34]作为检测器,基于少量样本得到初始模型,之后由初始模型为数据集产生伪标签,再借助更新后的数据集训练模型.随着自主学习的逐步深入,模型生成的样本逐渐变难,模型也变得更加鲁棒.由于数据集中各类标注数据极少,简单的自主学习极易陷入局部最小值.文献^[33]考虑多模型检测结果互补,将多模型学习整合到自主学习过程中,不同的模型通过设计正则项相互通信,进一步提高了所选训练样本的质量.同时由于多个模型之间的知识互补,有效提高了模型的精度和召回率.

3) 基于度量学习的方法

基于度量学习的小样本学习方法在图像识别任务中表现良好.一些研究者借鉴经典度量学习的思路解决目标检测问题同样取得了不错的效果.Zhang等^[35]提出单样本条件目标检测框架,名为比较网络.首先使用孪生网络分别提取查询图像和目标图像的特征作为边际概率在特征空间的度量.之后,引入包含可学习度量模块的两阶段检测器用于比较查询与目标图像特征进而得到检测结果.检测器包含条件候选区域网络和条件分类网络,两者均基于可学习度量模块获得条件概率估计.Karlinsky^[36]提出一种新的针对距离度量学习的方法,可以同时学习躯干网络参数、空间嵌入以及各类训练数据在该空间的多模态分布.基于此,他们提出一个子网络结构,并分别应用到小样本目标检测与小样本图像分类任务中,均取得了良好的效果.Kang等^[37]研究发现在具有丰富标注数据的基本类上训练目标检测模型时,模型可在网络顶层学习特定于某些目标属性的中间特征.这些特征能够隐式地组成只有极少数据的新类的高级表示.基于此,他们提出了能够从基础类别

调整中间特征并相应地检测新类别目标的模型.模型由基本检测模型和元模型组成.基础检测模型在基本类上进行训练用于提供基础特征,元模型学习基础特征用于输入图像的重要权重系数,进而通过调整基础特征重要程度帮助模型在小样本上进行检测.

3 目前小样本目标检测存在的问题

小样本目标检测作为极具现实意义的研究领域逐渐受到学术界与工业界的关注.但该领域的工作仍然较少,具有极大的发展空间.目前小样本目标检测主要存在以下几方面的问题.首先目前已有的工作大部分借助传统目标检测领域经典成熟的方法,如SSD、Faster-RCNN,采用两阶段的方式完成检测,使得模型部署后推理速度缓慢,大多不能满足实时要求.其次,基于候选框的检测方法虽然有效,但同时增加了模型结构的复杂度,导致在模型部署阶段对设备内存有了更高的需求.再者,现有方法的泛化性能差强人意,虽满足未见类别的设定,但是相近类别效果明显更优,随着小样本类别的增加,效果下降明显.另外,现有的方法大多是小样本学习与传统目标检测的简单融合,期待未来能有更加优秀的解决方案.最后,该领域仍未出现权威且普适的评测手段及评测标准.公平统一的验证结果有助于研究工作的比对与进步.

4 总结

通过分析现有的小样本目标检测方法,发现源知识到目标知识的迁移泛化、目标候选区域的提取、目标的分类与定位是小样本目标检测的关键.迁移学习中的正则化设置能够有效防止模型陷入过拟合,更好地进行知识迁移.借鉴注意力机制和采用多阶段检测方法,往往可以得到更准确的候选区域结果.借鉴度量学习的方法,在小样本目标检测中对于目标的精确分类有很大帮助.在未来的小样本目标检测研究中,可以尝试更多的元学习的方法,提高模型在多重任务多重场景下的泛化能力.

参考文献

References

- [1] Kuznetsova A, Rom H, Alldrin N, et al. The open images dataset v4: unified image classification, object detection, and visual relationship detection at scale[J]. arXiv Preprint, 2018, arXiv:1811.00982

- [2] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3) : 211-252
- [3] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: common objects in context [C] // European Conference on Computer Vision, 2014: 740-755
- [4] Koch G, Zemel R, Salakhutdinov R. Siamese neural networks for one-shot image recognition [C] // Proceedings of the 32nd International Conference on Machine Learning, 2015
- [5] Vinyals O, Blundell C, Lillicrap T, et al. Matching networks for one shot learning [C] // Advances in Neural Information Processing Systems, 2016: 3630-3638
- [6] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning [C] // Advances in Neural Information Processing Systems, 2017: 4077-4087
- [7] Sung F, Yang Y X, Zhang L, et al. Learning to compare: relation network for few-shot learning [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 1199-1208
- [8] Garcia V, Bruna J. Few-shot learning with graph neural networks [J]. arXiv Preprint, 2017, arXiv: 1711.04043
- [9] Kim J, Kim T, Kim S, et al. Edge-labeling graph neural network for few-shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 11-20
- [10] Santoro A, Bartunov S, Botvinick M, et al. One-shot learning with memory-augmented neural networks [J]. arXiv Preprint, 2016, arXiv: 1605.06065
- [11] Munkhdalai T, Yu H. Meta networks [C] // Proceedings of the 34th International Conference on Machine Learning-Volume 70, 2017: 2554-2563
- [12] Jamal M A, Qi G J. Task agnostic meta-learning for few-shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 11719-11727
- [13] Sun Q R, Liu Y Y, Chua T S, et al. Meta-transfer learning for few-shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 403-412
- [14] Ravi S, Larochelle H. Optimization as a model for few-shot learning [C] // 5th International Conference on Learning Representations, 2017
- [15] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks [C] // Proceedings of the 34th International Conference on Machine Learning (Volume 70), 2017: 1126-1135
- [16] Alfassy A, Karlinsky L, Aides A, et al. LaSO: label-set operations networks for multi-label few-shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 6548-6557
- [17] Zhang H G, Zhang J, Koniusz P. Few-shot learning via saliency-guided hallucination of samples [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 2770-2779
- [18] Chu W H, Li Y J, Chang J C, et al. Spot and learn: a maximum-entropy patch sampler for few-shot image classification [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 6251-6260
- [19] Chen Z T, Fu Y W, Wang Y X, et al. Image deformation meta-networks for one-shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 8680-8689
- [20] Zhang Z, He T, Zhang H, et al. Bag of freebies for training object detection neural networks [J]. arXiv Preprint, 2019, arXiv: 1902.04103
- [21] Schwartz E, Karlinsky L, Feris R, et al. Baby steps towards few-shot learning with multiple semantics [J]. arXiv Preprint, 2019, arXiv: 1906.01905
- [22] Schonfeld E, Ebrahimi S, Sinha S, et al. Generalized zero- and few-shot learning via aligned variational autoencoders [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 8247-8255
- [23] Wang X, Yu F, Wang R, et al. TAFE-net: task-aware feature embeddings for low shot learning [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 1831-1840
- [24] Li A X, Luo T G, Lu Z W, et al. Large-scale few-shot learning: knowledge transfer with class hierarchy [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 7212-7220
- [25] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [C] // European Conference on Computer Vision, 2016: 21-37
- [26] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788
- [27] 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
- [28] Donahue J, Jia Y Q, Vinyals O, et al. Decaf: a deep convolutional activation feature for generic visual recognition [C] // International Conference on Machine learning, 2014: 647-655
- [29] Razavian A S, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014: 806-813
- [30] Chen H, Wang Y L, Wang G Y, et al. LSTD: a low-shot transfer detector for object detection [J]. arXiv Preprint, 2018, arXiv: 1803.01529
- [31] Singh P, Varadarajan S, Singh A N, et al. Multidomain document layout understanding using few shot object detection [J]. arXiv Preprint, 2018, arXiv: 1808.07330
- [32] Fan Q, Zhuo W, Tai Y W. Few-shot object detection with attention-RPN and multi-relation detector [J]. arXiv Preprint, 2019, arXiv: 1908.01998
- [33] Dong X Y, Zheng L, Ma F, et al. Few-example object detection with model communication [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41 (7) : 1641-1654
- [34] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks [C] // Advances

- in Neural Information Processing Systems,2016:379-387
- [35] Zhang T F,Zhang Y,Sun X,et al.Comparison network for one-shot conditional object detection[J].arXiv Preprint,2019,arXiv:1904.02317
- [36] Karlinsky L,Shtok J,Harary S,et al.RepMet:representative-based metric learning for classification and few-shot object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2019:5197-5206
- [37] Kang B Y,Liu Z,Wang X,et al.Few-shot object detection via feature reweighting[J].arXiv Preprint,2018,arXiv:1812.01866

A survey of few-shot object detection

PAN Xingjia^{1,2} ZHANG Xulong^{1,2} DONG Weiming¹ YAO Hanxing³ XU Changsheng^{1,2}

1 Institute of Automation,Chinese Academy of Sciences,Beijing 100190

2 University of Chinese Academy of Sciences,Beijing 100190

3 Beijing LLVision Technology Co.,Ltd.,Beijing 100014

Abstract Research on object detection has developed rapidly in recent years with the progress of deep learning. However,the deep learning based object detection systems rely heavily on large scale labeled training data,which are rarely available in our realistic scene,so few-shot object detection get researchers' great concern.In this paper,we present a survey of few-shot object detection,and introduce the mainstream approaches and their characteristics,merits as well as limits.Finally,we provide the possible direction for further few-shot object detection research.

Key words few-shot learning;deep learning;object detection