

• 物联网 •

DOI:10.15961/j.jsuese.202100925



本刊网刊

基于知识蒸馏的目标检测模型增量深度学习方法

方维维, 陈爱方*, 孟 娜, 程虎威, 王清立

(北京交通大学 计算机与信息技术学院, 北京 100044)

摘 要: 随着万物互联时代的到来, 具备目标检测能力的物联网设备数量呈爆炸式增长。基于此, 网络边缘产生了海量的实时数据, 具有低时延、低带宽成本和高安全性特点的边缘计算随之成为一种新兴的计算模式。传统的深度学习方法通常假定在模型训练前所有数据已完全具备, 然而实际的边缘计算场景中大量的新数据及类别往往随时间逐渐产生和获得。为了在训练数据成批积累和更新的条件下在资源有限的边缘设备上高效率地完成目标检测任务, 本文提出了基于多中间层知识蒸馏的增量学习方法 (incremental learning method based on knowledge distillation of multiple intermediate layers, ILMIL)。首先, 为了能够适当地保留原有数据中的知识, 提出了包含多个网络中间层知识的蒸馏指标 (multi-layer feature map RPN and RCN knowledge, MFRRK)。ILMIL 将教师模型和学生模型的中间层特征的差异加入模型训练, 相比于现有的基于知识蒸馏方法的增量学习, 采用 ILMIL 方法训练的学生模型可以从教师模型的中间层学习到更多的旧类信息来缓解遗忘。其次, ILMIL 利用 MFRRK 蒸馏知识完成现有模型的增量训练, 避免训练使用多个独立模型带来的资源开销; 为进一步降低模型复杂度以高效地在边缘设备上部署推理, 可在知识蒸馏前进行剪枝操作来压缩现有模型。在不同场景和条件下的实验对比表明, 本文方法可在有效降低模型计算和存储开销的前提下, 缓解已有知识的灾难性遗忘现象, 并维持可接受的推理精度。

关键词: 目标检测; 边缘计算; 增量学习; 知识蒸馏; 模型压缩

中图分类号: TP391.41

文献标志码: A

文章编号: 2096-3246(2022)06-0059-08

Incremental Deep Learning Method for Object Detection Model Based on Knowledge Distillation

FANG Weiwei, CHEN Aifang*, MENG Na, CHENG Huwei, WANG Qingli

(School of Computer and Info. Technol., Beijing Jiaotong Univ., Beijing 100044, China)

Abstract: With the advent of the Internet of Everything era, the number of IoT devices with object detection capability has exploded. Accordingly, massive amounts of real-time data are generated at the edges of the network. Thus, edge computing has become an emerging computing paradigm that has the characteristics of low latency, low bandwidth and high security. While the traditional deep learning approaches usually assume that all data have been obtained before model training, a large number of new data and new categories are often gradually generated and obtained over time in real edge computing environment. In order to execute the object detection task efficiently on resource-constrained edge devices when the input data samples are accumulated and updated in batches, an incremental learning method based on knowledge distillation of multiple intermediate layers (ILMIL) was proposed in this paper. First, to preserve the obtained knowledge from existing data, a new metric called MFRRK was proposed for covering knowledge from multiple intermediate network layers. The discrepancy between the intermediate layers' features of teacher model and student model were added in ILMIL to model training. Compared to the recent incremental learning methods based on knowledge distillation, the student model adapted ILMIL was able to alleviate forgetting by learning more knowledge of the old classes from the intermediate layers of the teacher model. Then, the incremental training for current model was conducted to avoid resource costs of using mul-

收稿日期: 2021-09-14

基金项目: 国家自然科学基金项目 (62172031); 北京市自然科学基金-丰台轨道交通前沿研究联合基金项目 (L191019)

作者简介: 方维维 (1981—), 男, 副教授, 博士。研究方向: 物联网; 边缘计算与边缘智能。E-mail: fangww@bjtu.edu.cn

* 通信作者: 陈爱方, E-mail: 19125157@bjtu.edu.cn

网络出版时间: 2022-07-26 10:40:32

网络出版地址: <https://kns.cnki.net/kcms/detail/51.1773.TB.20220725.0954.013.html>

multiple independent models for training. To further reduce the model complexity, the model pruning technique can be used before knowledge distillation to compress the current model. Extensive experiments under different scenarios and conditions demonstrated that the proposed ILMIL method can effectively reduce the model calculation and storage costs, alleviate the catastrophic forgetting of existing knowledge, and maintain acceptable inference accuracy.

Key words: object detection; edge computing; incremental learning; knowledge distillation; model compression

随着物联网及4G/5G技术的发展,位于网络边缘的设备数量呈爆炸式增长趋势。同时,网络边缘产生的数据量也在急剧增加。传统云计算模式需要通过网络将大量数据发送到云计算中心进行处理,从而导致网络传输带宽的负载量增加,进而造成推理时延增加,因此无法完成自动驾驶、增强现实等需要获取实时检测结果的计算任务。由于传统云计算方式存在以上问题,数据在网络边缘设备上进行处理,边缘计算新模式具有低延迟、低带宽和高安全性等特点^[1-2],更适用于满足目标检测应用需求。

最早提出的目标检测模型如Viola-Jones(VJ)检测器^[3]等均是基于手工特征实现。随着手工设计特征的模型性能趋于饱和,基于深度学习的目标检测RCNN^[4]、Fast R-CNN^[5]和Faster R-CNN^[6]开始出现,其中Faster R-CNN以端到端的方式完成推理。为了加速模型推理,YOLO^[7]系列算法被提出。为了获得精度和速度的平衡,SSD^[8]算法被提出。为取消锚框的使用,近年来研究人员提出了基于关键点检测的算法CornerNet和CenterNet等^[9]。传统的目标检测之类的深度学习方法通常建立在模型训练前所有数据均已完全获得的前提条件下,然而在实际边缘计算应用中,新的数据样本和新的数据类别往往是逐渐产生、积累和获得的。所以,传统方法都会面临一个挑战——灾难性遗忘,即当接收到新数据时,模型会趋向于拟合新数据而欠拟合旧数据。

为了解决上述挑战,越来越多的研究开始关注增量学习(incremental learning, IL)^[10]这一主题。增量学习方法能够不断地从新样本中学习到新知识,同时也会保存大部分之前已学习的旧知识。目前的增量学习方法主要包括基于参数回放、基于动态网络和基于正则化3大类。基于参数回放的方法主要思想是从旧类数据集中挑选出重要数据进行保存,利用保存的部分旧类数据和所有新类数据完成模型的训练,例如:iCaRL^[10]中先使用特征均值衡量旧类数据的重要性,再保存靠近特征均值的旧类数据;Wang^[11]、Kemker^[12]等在宽度或深度维度上扩大模型或者训练多个模型,来增加模型的表示能力,以避免模型在新数据上的过拟合;为这些新的数据和类别独立地训练新的模型固然可行,但模型的部署将会给资源受限的边缘设备带来非常大的计算资源开销

和管理维护成本。基于动态网络的方法为每个类的执行固定不同的前馈通路,例如, PackNet^[13]算法中利用权重剪枝技术为每个类挑选出最优的前馈通路,但是需要进行多次剪枝和重训练,导致计算成本增加。基于正则化的方法主要对神经网络的权值更新施加约束,例如:LwF^[14]、LwM^[15]中使用知识蒸馏技术延缓旧类知识的遗忘;Chen等^[16]将增量学习方法用于目标检测领域。此类基于正则化的方法采用的知识蒸馏技术中,学生模型只学习了教师模型最后的logits知识,而教师模型的中间层特征图的知识大部分无法学习到,这使得大模型无法部署或部署在边缘设备上无法获得好的性能。

因此,针对上述问题及输入数据样本成批更新替换且边缘设备硬件资源受限的应用场景,本文提出了一种高效的基于多中间层知识蒸馏的增量学习方法ILMIL。首先,ILMIL将在旧数据上训练好的模型权重迁移到教师网络,并将压缩后的模型作为学生网络。其次,教师网络通过设计的存储多个网络中间层知识的MFRRK蒸馏指标来训练学生网络,最终得到一个存储友好的高性能学生模型。该方法可在有效降低模型计算和存储开销的前提下,缓解已有知识的灾难性遗忘现象,并维持可接受的推理精度。

1 ILMIL方法

1.1 总体流程

本文使用的目标检测模型是基于VGG16^[17]的Faster R-CNN,原因如下:1)实现了端到端的检测流程,在两阶段目标检测模型中具有较高的检测速度;2)相比于YOLO和SSD等常用模型,具有较高的检测精度;3)其区域建议网络(region proposal network, RPN)设计包含有目标定位信息,有助于缓解增量学习过程中旧类知识的遗忘。

图1展示了本文提出的ILMIL方法的总体结构。本文采用“教师-学生(T-S)”知识蒸馏架构^[18],数据集划分为旧类 D_o 和新类 D_n ,其中,旧类数据集的标签中仅含有旧类目标 C_o ,新类数据集的标签中仅含有新类目标 C_n ,且 $C_o \cap C_n = \emptyset$ 。使用 D_o 训练得到教师模型,将其复制一份得到初始的学生模型。在增量学习过程中,使用 D_n 训练学生模型,同时,教师模型的指导降低其对旧知识的遗忘速度,最终使得学生模型

可以识别所有的目标 $C = C_o \cup C_n$ 。如图1所示,一般情况下该增量学习过程可通过步骤①、②和⑤完成。如考虑到边缘设备的资源,对目标检测模型的执行速

度和模型大小有进一步的要求,则可选择在获得教师模型后进行适度的剪枝压缩,即通过步骤①、③、④和⑤完成。

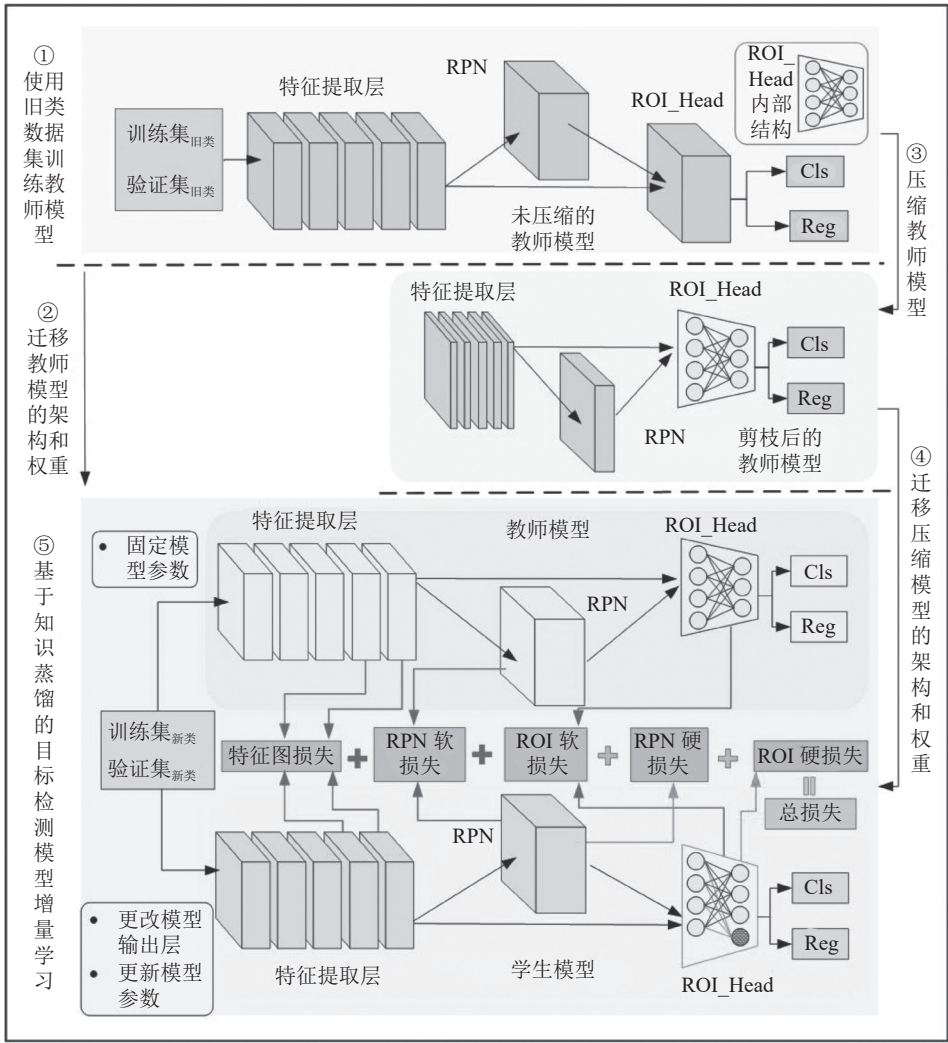


图1 ILMIL方法概述

Fig. 1 Overview of the ILMIL approach

1.2 知识蒸馏指标和方法设计

1.2.1 知识蒸馏指标

本文提出一种新的知识蒸馏指标MFRK (multi-layer feature map, RPN and RCN knowledge), 其融合了多个中间层的特征图信息、RPN, 以及分类和回归网络 (classification and regression network, RCN)。在现有的文献[17]研究中, 将Faster R-CNN特征提取模块VGG16的卷积层部分划分为Stage1、Stage2、Stage3、Stage4和Stage5。在增量学习中, 教师模型中间层的特征图包含了大量的旧类信息, 通过最小化学生模型和教师模型特征图之间的差异, 可帮助学生模型有效地保留旧类知识, 因此, 将特征提取模块的特征图信息作为蒸馏知识所考虑的指标之一。对于不同深度的特征图, 提取出的信息侧重

点不同, 其中: 越靠近图片输入层的特征图越关注具体的图像细节特征, 比如纹理特征; 越远离输入层的特征图越关注图像的语义信息, 即抽象信息。所以, 与传统相关知识蒸馏技术^[14,16]不同的是, 本文在选取该部分指标时, 除选取Stage5模块的输出特征图外, 还选取了Stage4模块的输出特征图。通过向教师模型学习不同层次的特征图知识, 使学生模型可以学习到旧类目标的细节信息和语义信息, 从而缓解学生模型对旧类知识的遗忘现象。与现有的利用知识蒸馏完成增量学习方法^[13,19]相同的是, 本文采用的知识蒸馏方法也使用了RPN和RCN的输出信息, 其中: RPN包含了目标建议; RCN针对每一个目标建议, 再给出类别评分和空间位置调整向量。

1.2.2 知识蒸馏方法

将参与增量训练的数据集合表示为 $\{x, y, y_{\text{reg}}\}$, 其中, x 为输入图像, y 为目标所属的类别, y_{reg} 为目标的真实位置信息。教师模型的输出表示为 $(P_T^{\text{RPN}}, R_T^{\text{RPN}}, R_T^{\text{RCN}})$, 学生模型的输出表示为 $(P_S^{\text{RPN}}, R_S^{\text{RPN}}, P_S^{\text{RCN}}, R_S^{\text{RCN}})$, 其中, $P_T^{\text{RPN}}, R_T^{\text{RPN}}, P_T^{\text{RCN}}$ 和 R_T^{RCN} 分别为教师模型的RPN模块、RCN模块输出的分类与定位信息, $P_S^{\text{RPN}}, R_S^{\text{RPN}}, P_S^{\text{RCN}}$ 和 R_S^{RCN} 分别为学生模型的RPN模块、RCN模块输出的分类与定位信息。以上在教师模型和学生模型中涉及到的所有分类输出 P 表示为 $P = \text{softmax}(Z/T)$, 其中 Z 为模型最终的输出分数。

在对学生模型进行增量训练的过程中使用的总损失函数包括硬损失 L_{hard} 、软损失 L_{soft} 和提示损失 L_{hint} , 总损失函数如式(1)所示:

$$L_{\text{total}} = L_{\text{hard}} + L_{\text{soft}} + \gamma L_{\text{hint}} \quad (1)$$

式中: 硬损失 L_{hard} 由RPN和RCN的分类损失 $L_{\text{hard-clc}}^{\text{RPN}}, L_{\text{hard-clc}}^{\text{RCN}}$, 以及定位损失 $L_{\text{hard-reg}}^{\text{RPN}}, L_{\text{hard-reg}}^{\text{RCN}}$ 组成; γ 为控制不同损失之间平衡的超参数, $\gamma > 0$; 软损失 L_{soft} 由RPN、RCN和学生模型的输出信息设计而成; 中间层的提示损失 L_{hint} 由差异构成。 $L_{\text{hard}}, L_{\text{soft}}, L_{\text{hint}}$ 的表示分别如式(2)、(3)、(4)所示:

$$L_{\text{hard}} = L_{\text{hard}}^{\text{RPN}} + L_{\text{hard}}^{\text{RCN}} = \left(L_{\text{hard-clc}}^{\text{RPN}}(P_S^{\text{RPN}}, y) + L_{\text{hard-reg}}^{\text{RPN}}(R_S^{\text{RPN}}, y_{\text{reg}}) \right) + \left(L_{\text{hard-clc}}^{\text{RCN}}(P_S^{\text{RCN}}, y) + L_{\text{hard-reg}}^{\text{RCN}}(R_S^{\text{RCN}}, y_{\text{reg}}) \right) \quad (2)$$

$$L_{\text{soft}} = L_{\text{soft}}^{\text{RPN}} + L_{\text{soft}}^{\text{RCN}} = \left(L_{\text{soft-clc}}^{\text{RPN}}(P_S^{\text{RPN}}, P_T^{\text{RPN}}) + L_{\text{soft-reg}}^{\text{RPN}}(R_S^{\text{RPN}}, R_T^{\text{RPN}}) \right) + \left(L_{\text{soft-clc}}^{\text{RCN}}(P_S^{\text{RCN}}, P_T^{\text{RCN}}) + L_{\text{soft-reg}}^{\text{RCN}}(R_S^{\text{RCN}}, R_T^{\text{RCN}}) \right) \quad (3)$$

式中, $L_{\text{soft-clc}}^{\text{RPN}}, L_{\text{soft-clc}}^{\text{RCN}}, L_{\text{soft-reg}}^{\text{RPN}}, L_{\text{soft-reg}}^{\text{RCN}}$ 分别为RPN和RCN的分类和定位软损失。

$$L_{\text{hint}} = \gamma[(T_4 - S_4)^2 + (T_5 - S_5)^2]/2 \quad (4)$$

式中, T_4, T_5 和 S_4, S_5 分别为教师网络 T 和学生网络 S 的Stage4和Stage5的输出特征图。

式(2)和(3)中: 令*号代表soft和hard, 分类损失 $L_{*-\text{cls}}$ 均为交叉熵损失, 如式(5)所示; 回归损失 $L_{*-\text{reg}}$ 均为平滑L1损失, 如式(6)所示。

$$L_{*-\text{cls}}(p, q) = - \sum_{i=1}^m p(x_i) \text{lb}(q(x_i)) \quad (5)$$

$$L_{*-\text{reg}}(t, v) = \sum_{i=1}^n \text{smooth}_{L_1}(t_i - v_i) \quad (6)$$

式中, p 和 q 分别为两个离散概率分布, t 和 v 分别为两组不同的目标框位置标注, m 和 n 分别为类别和目标框的总数, $\text{smooth}_{L_1}()$ 表达式为:

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1; \\ |x| - 0.5, & \text{其他} \end{cases}$$

1.3 基于知识蒸馏的模型训练流程

如图1所示, 为了在原有的旧类知识基础上融入新的知识, 同时避免灾难性遗忘, 基于第1.2节中定义的指标和方法进行增量模型训练, 流程如下:

1) 训练教师模型 T 。与旧类数据集相对应, 教师模型的输出层神经元个数为 $|C_0|+1$, 其中, $|C_0|$ 为旧类目标的个数, “1” 为背景类。将训练完毕后的以上模型作为教师模型 T 和初始的学生模型。

2) 构建学生模型 S 。将拷贝得到的学生模型输出层神经元个数修改为 $C+1$, 其中, C 为总目标数量, “1” 表示背景类。随机初始化输出层中新添加神经元的权重。修改后的学生模型 S 将参与蒸馏训练。

3) 增量训练。基于1)中的 T 和2)中的 S 构成 $T-S$ 知识蒸馏架构, 进行第1.2.2节中提出的蒸馏方法完成增量模型训练。

1.4 结合剪枝压缩的模型增量训练流程

通过第1.3节所述增量训练流程所得到的学生模型的总体计算和存储开销与原教师模型几乎相同, 避免了独立训练和使用多个检测不同种类目标的模型所带来的较高资源开销。但是, 边缘侧设备往往资源较为受限。为有效降低推理过程中的计算和存储开销, 可在知识蒸馏前对教师模型执行基于剪枝技术的压缩操作, 即图1中的步骤③。

本文在模型压缩阶段使用的是基于L1范数的通道剪枝算法^[20], 主要原因如下: 1) 通道剪枝属于结构化剪枝, 其实现不依赖于任何稀疏卷积计算库和专门设计的硬件; 2) 实现较为简单, 表征通道重要性的L1范数值可以直接通过计算获得, 无需额外操作; 3) 该剪枝算法可以在压缩的同时, 有效地维持原模型的推理精度^[20]。

2 实验结果分析

2.1 实验环境

使用Python 3.6.9开发语言在PyTorch 1.1.0深度学习框架上进行模型构建和实验, 选择PASCAL VOC 2007作为数据集, 该数据集共包含了20个类别的图像数据。训练过程中, 使用了随机梯度下降法(SGD)作为反向传播算法, SGD的动量为0.9, 初始学习率为0.000 1。在对未压缩模型的增量训练中, 共迭代20轮次, 其中, 每迭代10轮次, 学习率下降为当前的0.1倍; 在对压缩后模型的增量训练中, 共迭代30轮次, 其中, 每迭代5轮次, 学习率下降为当前的0.5倍。本文采用mAP(mean average precision)^[21]衡量模型准确度, 采用FLOPS(floating point operations per

second) 衡量模型的计算量。

2.2 对未压缩模型的增量训练

为了验证本文方法在新类出现时的有效性, 设置了不同的实验, 其中涉及到的实验组别及训练方法含义如表1所示。

表 1 实验组别	
Tab. 1 Experiment groups	
训练方法	表示含义
A(1~20)	用所有20个类训练教师网络, 不做增量学习, 作为对比参考
A(1~19)	用前19个类训练教师网络, 不做增量学习
A(1~10)	用前10个类训练教师网络, 不做增量学习
+B(20)	用前19个类训练教师网络, 再加入第20个类进行训练
+B(11~20)	用前10个类训练教师网络, 再加入剩下的10个类进行训练
+D(16)	用前15个类训练教师网络, 再加入第16个类进行训练
+D(17)	用+D(16)得到的增量模型作为教师网络, 再加入第17个类进行训练
+D(18)	用+D(17)得到的增量模型作为教师网络, 再加入第18个类进行训练
+D(19)	用+D(18)得到的增量模型作为教师网络, 再加入第19个类进行训练
* c	蒸馏知识为RPN和RCN的分类信息
* c+b	蒸馏知识为RPN和RCN的分类和回归信息
* c+b+h5 ^[12]	蒸馏知识为RPN、RCN的分类和回归信息, Stage5的信息
本文	蒸馏知识为RPN、RCN的分类和回归信息, Stage4和Stage5的信息
* c+b+h4+h5	

注: * 表示一个任意叠加的标志符。

1) 单次增加一个类

为了验证本文提出的方法在单次增加一个新类别时的有效性, 一次性使用前19个类训练教师网络, 再加入第20个类进行训练。图2展示了未压缩模型单次增加一个类的mAP的对比实验结果。如图2所示: 对学生模型微调后, 旧类的mAP从70.8降到18.4。与其他指标和方法对比, 本文提出的增量学习方法在旧类别目标和总目标上均维持了最高的mAP, 降低了灾难性遗忘问题对旧类别的影响, 同时获得了较高的新类别mAP。

2) 单次增加多个类

图3展示了未压缩模型单次增加多个类的mAP的对比实验结果。如图3所示, 未进行知识蒸馏的训练中旧类的mAP从72.1降为7.2。本文提出的增量学习方法得到的实验结果在总目标的mAP结果上达到了最优, 同时在旧类和新类上的检测mAP也优于绝大多数方法, 验证本文提出的方法在单次增加多个新类别时的有效性。

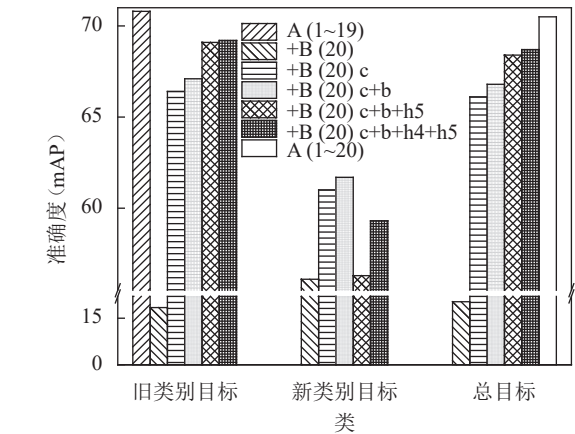


图 2 单次增加一个类的实验结果

Fig. 2 Experimental results of adding a class

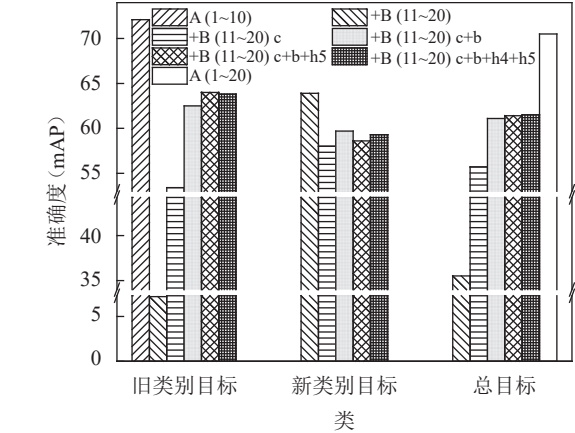


图 3 单次增加多个类的实验结果

Fig. 3 Experimental results of adding multiple classes

3) 逐步增加多个类

使用前一次训练 (如+D(16)) 得到模型作为当前训练的教师模型 (如+D(17))。表2展示了未压缩模型逐步增加多个类的mAP的对比实验结果。

表 2 逐步增加多个类的实验结果			
Tab. 2 Experimental results of gradually adding multiple classes			
训练方法	mAP		
	旧类别目标	新类别目标	总目标
+D(16)	17.1	29.4	17.9
+D(16) c+b+h5	70.5	27.2	67.8
本文+D(16) c+b+h4+h5	70.7	27.4	68.0
+D(17)	17.1	26.1	18.1
+D(17) c+b+h5	64.5	42.6	63.2
本文+D(17) c+b+h4+h5	65.2	39.3	63.7
+D(18)	9.8	28.2	10.8
+D(18) c+b+h5	62.5	33.7	60.9
本文+D(18) c+b+h4+h5	62.5	37.0	61.1
+D(19)	8.8	31.2	10.0
+D(19) c+b+h5	58.4	37.9	57.4
本文+D(19) c+b+h4+h5	59.5	42.3	58.6

如表2所示,直接训练会造成灾难性遗忘问题,旧类别目标的mAP从17.1快速降到8.8,总目标从17.9快速降到10.0。与其他方法相比,本文提出的增量学习方法在逐步增加多个类时,维持了相对较高的旧类、新类和总目标mAP,旧类的mAP在59.5~70.7之间,新类的mAP在27.4~42.3之间,而总目标mAP在58.6以上,这说明提出的ILMIL方法使学生模型额外学习到了教师模型的中间层Stage4、Stage5的特征知识,以及本文设计的目标函数所包含的硬损失、软损失和提示损失起到了关键作用。

4)超参数 γ 的影响

图4展示了超参数 γ 的不同取值对mAP的影响。 γ 作为超参数对式(1)中的提示损失和其他损失进行平衡,从而影响学生模型在各个类别上的检测mAP。以单次增加一个类为例进行实验。如图4所示:将 γ 设为1时,蒸馏训练既使学生模型没有忘记旧类的知识,也学习到了新类的信息;模型对旧类和总类的检测mAP最高,对新类的检测mAP居中。因此本文实验中均将该参数值设定为1。

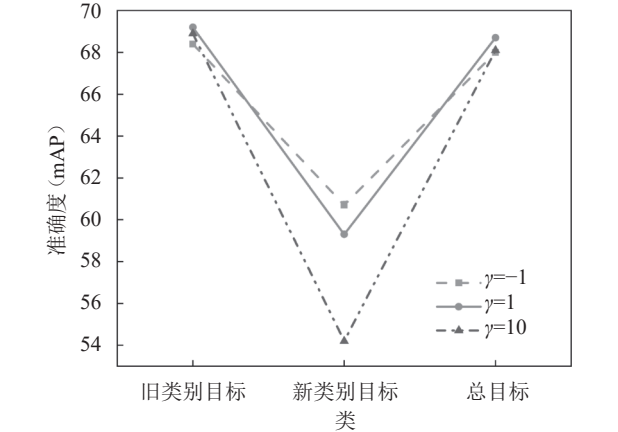


图 4 参数 γ 的影响

Fig. 4 Influence of parameter γ

2.3 对压缩后模型的增量训练

1)构建学生模型

构建了4个不同压缩率的模型,图5以单次增加一个类为例,展示了1个原始模型和4个压缩模型的数量和计算量。两个实验场景使用的原始模型和压缩模型的数量和计算量与图5中的数据一致。其中,GFLOPS(giga floating point operations per second)表示 10^9 FLOPS。

如图5所示:原始模型的数量为130.48 MB,计算量为728.32 GFLOPS。学生模型1的压缩比例为18.2%,其数量为106.8 MB,计算量为616.3 GFLOPS;学生模型2的压缩比例为29.9%;学生模型3的压缩比例为35.6%,数量和计算量逐步降低;学生模型4具有最

大的压缩比例46.8%,其数量为69.4 MB,计算量为406.8 GFLOPS。

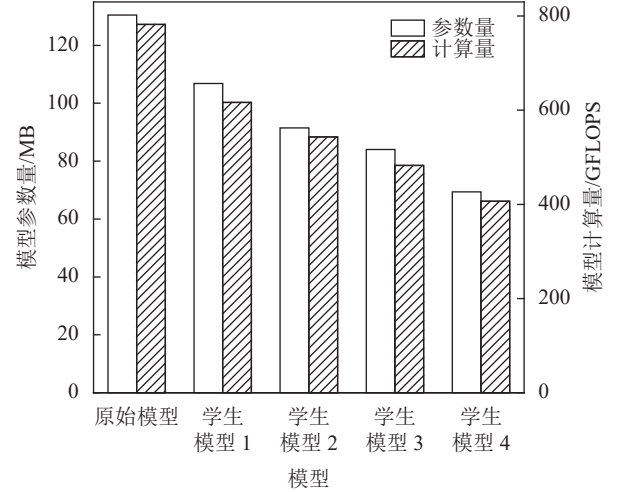


图 5 模型参数量和计算量

Fig. 5 Amount of model parameters and computation

2)单次增加一个类

图6展示了非压缩模型与4个学生模型对应的压缩率取值,以及在增加一个类后对应的mAP的实验结果。如图6所示:非压缩模型在旧类、新类和总类上的检测mAP均最高。随着压缩比例不断升高,压缩模型在各类上的检测精度均在缓慢降低,其中,以总类mAP为例,非压缩模型为68.7,最低压缩比例为18.2%的模型1为65.7,最高压缩比例为46.8%的模型4为59.4。总体来看,利用第1.4节介绍的增量训练流程既可有效降低目标检测模型的复杂度,也可有效缓解对旧类别知识的灾难性遗忘现象,在一定mAP损失内,采用的模型压缩方法极大地降低了模型的大小。

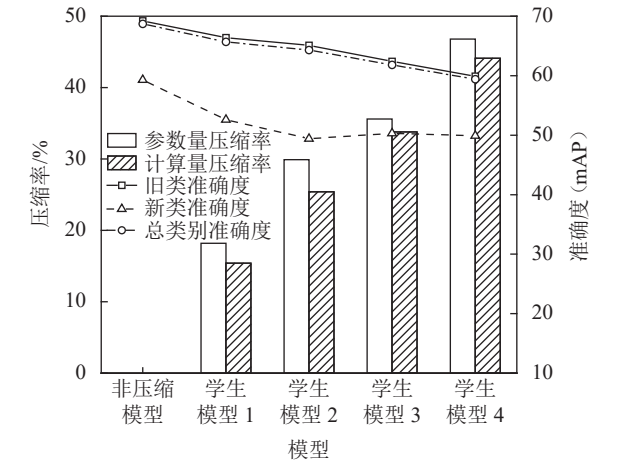


图 6 压缩后增加一个类的实验结果

Fig. 6 Experimental results of adding a class after compression

3)单次增加多个类

图7展示了非压缩模型与4个学生模型对应的压

缩率取值,以及在增加多个类后对应的mAP的实验结果。如图7所示:随着模型压缩比例的增大,在各类上的检测mAP总体呈下降趋势。从各模型对总类的检测mAP来看,非压缩模型为61.5,最低压缩比例为18.2%的模型1为61.3,最高压缩比例为46.8%的模型4为59.8。总体来看,在目标检测模型复杂度降低的同时,压缩后的模型的mAP损失降低在可接收范围内,只有0.2~1.7,且在增量过程中有效缓解了灾难性遗忘现象,说明本文所提出的增量学习方法在压缩率较大,模型参数量减少的情况下,依然能有效地学习到教师模型的中间层信息,保证了模型的性能。

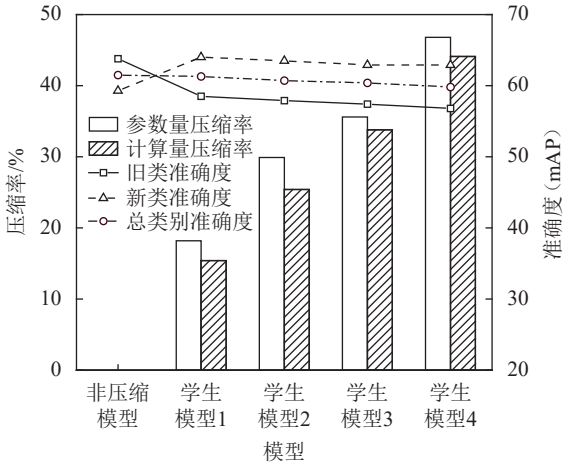


图7 压缩后增加多个类的实验结果

Fig. 7 Experimental results of adding multiple classes after compression

3 结论

在具备目标检测能力的物联网设备不断增加,从而导致网络边缘数据爆炸式增长的场景下,针对目标检测实际应用中深度学习所需的训练数据及其类别难以一次性完全获得的实际问题,考虑到网络边缘设备的资源限制,本文提出了一种基于知识蒸馏的目标检测模型增量深度学习方法。首先,提出了包含多个网络中间层知识的蒸馏指标;然后,在此基础上提出了相应的增量模型训练方法,该方法可进一步与模型剪枝压缩技术相结合来适应边缘设备的资源能力;最后,基于PASCAL VOC 2007目标检测数据集进行了在多种不同增量组合方式和模型训练方式下的对比实验。实验结果表明,本文所提出的增量学习方法能够有效地避免和缓解灾难性遗忘现象的发生,维持较高的目标检测性能。通过与剪枝技术相结合,可进一步减少模型参数大小和计算总量,便于在边缘设备上部署和使用。对于未来的研究工作:一个方向是,准备通过在更多的DNN模型和目标检测数据集上的评估来完善ILMIL的方法设计;另一个方

向是,采用ILMIL方法解决其他相关类型的任务,例如,在图像分类^[22]中结合内存高效的增量学习方法^[23]和推理高效的早期退出^[24]方法来解决遗忘问题,以及结合少样本学习来完成图像分割^[25]中的增量学习任务等等。

参考文献:

- [1] Cao Keyan, Liu Yefan, Meng Gongjie, et al. An overview on edge computing research[J]. *IEEE Access*, 2020, 8: 85714–85728.
- [2] Shi Weisong, Cao Jie, Zhang Quan, et al. Edge computing: Vision and challenges[J]. *IEEE Internet of Things Journal*, 2016, 3(5): 637–646.
- [3] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]// *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001. Kauai: IEEE, 2001: 1.
- [4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014: 580–587.
- [5] Girshick R. Fast R-CNN[C]// *Proceedings of the 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1440–1448.
- [6] Ren Shaoqing, He Kaiming, 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.
- [7] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 779–788.
- [8] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[M]// *European conference on computer vision*. Cham: Springer, 2016: 21–37.
- [9] Duan Kaiwen, Bai Song, Xie Lingxi, et al. CenterNet: Keypoint triplets for object detection[C]// *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019: 6568–6577.
- [10] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCaRL: Incremental classifier and representation learning[C]// *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 5533–5542.
- [11] Wang Yuxiong, Ramanan D, Hebert M. Growing a brain: Fine-tuning by increasing model capacity[C]// *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 3029–3038.
- [12] Kemker R, Kanan C. FearNet: Brain-inspired model for in-

- cremental learning[EB/OL].[2021-09-01].<https://arxiv.org/abs/1711.10563>.
- [13] Mallya A, Lazebnik S. PackNet: Adding multiple tasks to a single network by iterative pruning[C]// *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 7765–7773.
- [14] Li Zhizhong, Hoiem D. Learning without forgetting[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(12): 2935–2947.
- [15] Dhar P, Singh R V, Peng Kuanchuan, et al. Learning without memorizing[C]// *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019: 5133–5141.
- [16] Chen Li, Yu Chunyan, Chen Lvcai. A new knowledge distillation for incremental object detection[C]// *Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN)*. Budapest: IEEE, 2019: 1–7.
- [17] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL].[2021-09-01]. <https://arxiv.org/abs/1409.1556>.
- [18] Li Hao, Kadav A, Durdanovic I, et al. Pruning filters for efficient convnets[EB/OL]. (2016-08-31)[2021-09-01]. <https://arxiv.org/abs/1608.08710v1>.
- [19] Chen Guobin, Choi W G, Yu Xiang, et al. Learning efficient object detection models with knowledge distillation[C]// *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. New York: Curran Associates Inc., 2017: 742–751.
- [20] Courbariaux M, Bengio Y, David J P. BinaryConnect: Training deep neural networks with binary weights during propagation[C]// *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2015: 3123–3131.
- [21] Everingham M, Gool L, Williams C K I, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303–338.
- [22] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification[C]// *Proceedings of the 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1026–1034.
- [23] Zhao Hanbin, Wang Hui, Fu Yongjian, et al. Memory efficient class-incremental learning for image classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, PP(99): 1–12.
- [24] Teerapittayanon S, McDanel B, Kung H T. BranchyNet: Fast inference via early exiting from deep neural networks[C]// *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*. Cancun: IEEE, 2016: 2464–2469.
- [25] Ganea D A, Boom B, Poppe R. Incremental few-shot instance segmentation[C]// *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville: IEEE, 2021: 1185–1194.

(编辑 赵 婧)

引用格式: Fang Weiwei, Chen Aifang, Meng Na, et al. Incremental deep learning method for object detection model based on knowledge distillation[J]. *Advanced Engineering Sciences*, 2022, 54(6): 59–66. [方维维, 陈爱方, 孟娜, 等. 基于知识蒸馏的目标检测模型增量深度学习[J]. *工程科学与技术*, 2022, 54(6): 59–66.]