

doi:10.6041/j.issn.1000-1298.2021.08.022

基于非对称混洗卷积神经网络的苹果叶部病害分割

何自芬 黄俊璇 刘 强 张印辉
(昆明理工大学机电工程学院, 昆明 650500)

摘要: 针对苹果叶部病害由于数据集类间样本不均衡和拍摄角度、光照变化等实际成像与环境因素造成的精度低和泛化能力差的问题,本文提出了一种新型的非对称混洗卷积神经网络 ASNet。首先,通过在 ResNeXt 骨干网络中添加改进的 scSE 注意力机制模块增强网络提取的特征;其次,针对多数叶片病害特征分布相对分散的问题,使用非对称混洗卷积模块代替原始的残差模块来扩大卷积核的感受野和增强特征提取能力,从而提升模型的分割精度和泛化能力;最后,在非对称混洗卷积模块中使用通道压缩和通道混洗的方式弥补了分组卷积造成的通道间关联性不足的缺陷,降低了由于叶部病害类间不均衡导致的传统网络模型精度偏低的问题。在 COCO 数据集评价指标下,实验结果表明,相比于骨干网络为 ResNeXt-50 的原始 Mask R-CNN 模型,本文模型的平均分割精度达到 96.8%,提升了 5.2 个百分点,模型权重文件减小为 321 MB,减小了 170 MB。对实地采集和 AI Challenger 农作物病害分割挑战赛的 240 幅苹果叶片图像进行测试,结果表明,本文模型 ASNet 对苹果黑腐病、锈病与黑星病 3 种病害和健康叶片的平均分割精度达到 94.7%。

关键词: 苹果叶部; 病害分割; ASNet 模型; 非对称混洗卷积; 通道压缩; 注意力机制

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2021)08-0221-10 **OSID:** 

High Precision Identification of Apple Leaf Diseases Based on Asymmetric Shuffle Convolution

HE Zifen HUANG Junxuan LIU Qiang ZHANG Yinhui
(College of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Aiming at the problems of low accuracy and poor generalization ability caused by the imbalance of samples between data sets, shooting angles, light changes and other actual imaging and environmental factors caused by apple leaf diseases, a type of asymmetric shuffle convolution neural network ASNet was proposed. Firstly, by adding an improved scSE attention mechanism module to the ResNeXt backbone network to enhance the network feature extraction; secondly, for the relatively scattered feature distribution of most leaf diseases, the asymmetric shuffle convolution module was used to replace the original residual module to expand the receptive field of the convolution kernel and the enhanced feature extraction ability, thereby improving the recognition accuracy and generalization ability of the model; finally, the use of channel squeeze and channel shuffling in the asymmetric shuffle convolution module made up for the grouping convolution. The defect of insufficient correlation between channels reduced the problem of low recognition accuracy of traditional network models caused by the imbalance between leaf diseases. Under the COCO data set evaluation index, the experimental results showed that compared with the Mask R-CNN whose backbone network was ResNeXt-50, the average test accuracy of this model reached 96.8%, which was increased by 5.2 percentage points, and the model size was reduced to 321 MB, a decrease of 170 MB. Tested by 240 field-collected and AI Challenger crop disease identification challenge apple leaf images, the test results showed that the average segmentation accuracy of the proposed model ASNet for apple black rot, rust, scab and healthy leaves reached 94.7%.

Key words: apple leaf; disease segmentation; ASNet model; asymmetric shuffle convolution; channel squeeze; attention mechanism

收稿日期: 2021-01-15 修回日期: 2021-05-29
基金项目: 国家自然科学基金项目(61761024,62061022)
作者简介: 何自芬(1976—),女,副教授,博士,主要从事图像处理和机器视觉研究,E-mail: zyhhzf1998@163.com
通信作者: 张印辉(1977—),男,教授,博士,主要从事图像处理、机器视觉及机器智能研究,E-mail: yinhui_z@163.com

0 引言

苹果是我国产量最多的水果,同时中国也是世界上苹果产量最多的国家,占全世界的产量 50% 以上。苹果病害不仅影响苹果的产量和品质,也直接影响到国家和果农的经济收益,因此,苹果的病害识别技术对病害的防治很有必要^[1-3]。

传统的苹果病害类别往往依赖于专业人士的判断,这在很大程度上造成了果农在面对病害时不能采取相应的方法,造成大量经济损失。随着深度学习的兴起,农作物的病害叶片的分割和识别也逐渐成为研究热点。传统图像分割方法^[4-5]的主要思路是使用单一阈值将叶片的病害区域从背景中提取出来,再使用机器学习方法如支持向量机来对分割出的区域进行分类,但在自然条件下采集的数据集由于环境的影响(如光照、天气、背景等),实际成像结果与理想状况相差较大,因此单一阈值分割方法并不适用。同时,由于叶片病害的病斑特征分布较散,普通的阈值分割也很难将病害特征完全提取出来。在深度学习方面,卷积神经网络具有自动从图像中逐层提取特征的优点,因此深度学习在农作物病害分割和识别的应用上有了长足发展^[6]。文献[7]在 VGG-16 模型的基础上,设计了全新的全连接层模块,将 ImageNet 图像数据集训练的卷积层迁移到 VGG-16 模型中,并对仅训练全连接层的模型和训练全部层(卷积层+全连接层)的模型,两种迁移学习方式进行了试验,最终在扩充数据集和训练全部层的情况下,模型对玉米健康叶片和 2 种病害图像的平均识别精度达到 95.33%。文献[8-10]都提出了利用多尺度卷积融合的方法来增强网络提取的特征,从而获得了较好的测试效果,但在实地测试时精度相对较低。文献[11-12]提出通过设置不同卷积核函数和池化函数来分析模型最终的准确性,使得精度有一定提升。文献[13]提出了一种基于神经结构搜索的多种农作物的病害图像识别方法,并探究了图像在缺乏颜色信息时对该结构的影响。但是当输入数据为未经处理的类别不均衡图像时,精度下降较大,因此该模型的泛化能力较差。同时,该网络需要的数据量较为庞大,而现今的单类别农作物数据集,在数据量和类别均衡程度上均很难达到要求,PlantVillage 数据集中苹果锈病的病害图像仅 274 幅,为其他类别图像的 1/3 甚至更少,因此需要提前对数据集进行大量的扩充和增强。在实验数据方面,该模型训练和测试均采用背景单一、环境因素影响较小的公共数据集,并没有使用实地采集的数据集进行测试,因此难以确定该模型在实际应用

中的效果,增加了该模型的应用难度。文献[14]针对复杂场景下的农作物病害分割提出了一种融合多特征图切割的病斑自动提取方法,为了去除病害图像中的背景,在两层框架下引入了显著性检测法,同时提出了一种病斑袋模型对网络的中层特征进行有效提取,针对检索单一特征时分割率偏低的问题,采用自适应融合策略对病害图像全局和局部特征进行融合,但自然条件下的叶片病害图像的病斑存在边界不明确、遮挡、旋转等影响分割精度的因素,因此该分割方法并没有取得较高的精度,使得该模型对多种农作物病害分割的平均精度相对较低,分割正确率仅为 90.8%。

本文基于 Mask R-CNN 提出一种非对称混洗卷积神经网络(Asymmetric shuffle network, ASNet),该模型在保留 Mask R-CNN 优点的同时,通过在骨干网络 ResNeXt 中添加改进的 scSE 注意力机制以增强网络提取的特征;通过通道压缩和通道混洗方法降低数据集总量较少和类别不均衡对精度的影响;通过非对称混洗卷积的方式增强网络的泛化能力和特征提取能力。

1 模型选择

1.1 骨干网络选择

ResNeXt^[15]结合了 Inception^[16]、ResNet^[17]和分组卷积(Group convolution, GConv)^[18]的优点,在保留了残差结构的同时,将原来的 Block 结构分解成多个相同的分支结构,输入图像被分组送入各个分支结构中分别进行卷积操作,将所有分支输出的特征图进行维度拼接得到最终的输出结果。由于每个分支结构都是由 2 组 1×1 和 1 组 3×3 卷积层堆叠而成的相同结构,因此网络总体参数量相较于 ResNet 骨干网络并没有大量增加。

1.2 Mask R-CNN 网络模型

单独的检测网络分别对目标进行了定位和类别预测,即使目标分类的置信度较高,依然存在一定程度的定位不精确问题。Mask R-CNN^[19]作为增加了掩码(Mask)预测分支的多任务训练网络,既保证了较高的目标分类精度,同时像素级的分割又极大提升了目标的定位精度。

Mask R-CNN 在 Faster R-CNN 的基础上增加了 Mask 预测分支,由原来的 2 个任务(分类+回归)增加到 3 个任务(分类+回归+分割),而多任务训练网络能增强模型的泛化能力,同时提高网络的精度。Mask R-CNN 将二进制 Mask 与来自 Faster R-CNN 的分类和边界框组合,产生了精确分割的图像,实现了 Mask 和类别预测关系的解耦,对

每一个目标物体,不仅给出其边界框,并且对边界框内的各个像素是否属于该物体进行了标记。所以 Mask R - CNN 是灵活通用的对象实例分割框架,可以完成目标分类、目标检测、语义分割、实例分割和人体姿势分割等多种任务。

1.3 Mask Scoring R - CNN 网络模型

Mask Scoring R - CNN^[20]是在 Mask R - CNN 基础上改进得到的。在实例分割中,Mask 得分都是由实例分类的置信度分数决定。由于 Mask 的质量被量化为 Mask 实例与其真实值 (Ground truth, GT) 之间的 IoU,因此 Mask 的得分与分类置信度的相关性不强。针对这个问题,Mask Scoring R - CNN 设计了一种新的分支结构 MaskIoU Head。该结构将实例特征与相应的 Mask 结合起来,对 MaskIoU 进行回归,矫正 Mask 质量和 Mask 得分之间的偏差,使网络在训练过程中对 COCO 数据集的评估优先考虑更加准确的 Mask 预测,从而提高实例分割的精度。

Mask Scoring R - CNN 网络模型在 COCO 数据集上取得了较好的结果,但在面对小样本的数据集时,Mask Scoring R - CNN 网络结构较为复杂,参数存在大量冗余,时空复杂度较高。并且在本文数据集中,苹果叶部病害特征的早期颜色相似度较高,病

害的病斑在叶片上空间分布广泛,模型使用的 3×3 卷积核的感受野较小,无法准确提取病害的特征,因此 Mask Scoring R - CNN 网络模型分割错误率增高,精度损失较大。

针对上述问题,本文以增强网络特征提取能力作为切入点,使用非对称混洗卷积结构的 3 层并行分支共同提取图像特征并将结果进行逐像素相加以丰富目标的特征空间,提升模型精度;同时对 3 层分支结构分别使用分组卷积或深度可分离卷积以减少网络模型的参数量,降低时空复杂度。

2 非对称混洗卷积神经网络

2.1 ASNet 网络结构

为实现小样本数据集的苹果叶部病害的高精度分割,本文提出一种精度高、泛化能力强的神经网络 ASNet,网络结构如图 1 所示。改进方法如下:

- (1) 针对数据集样本较小、目标类别不均衡的问题,提出了通道压缩和通道混洗的方法,降低了网络模型对输入图像数量的依赖性。
- (2) 提出了非对称混洗卷积代替 3×3 卷积以增强网络的特征提取能力和泛化能力。
- (3) 改进注意力机制,扩大重要特征信息权重以提高模型的精度。

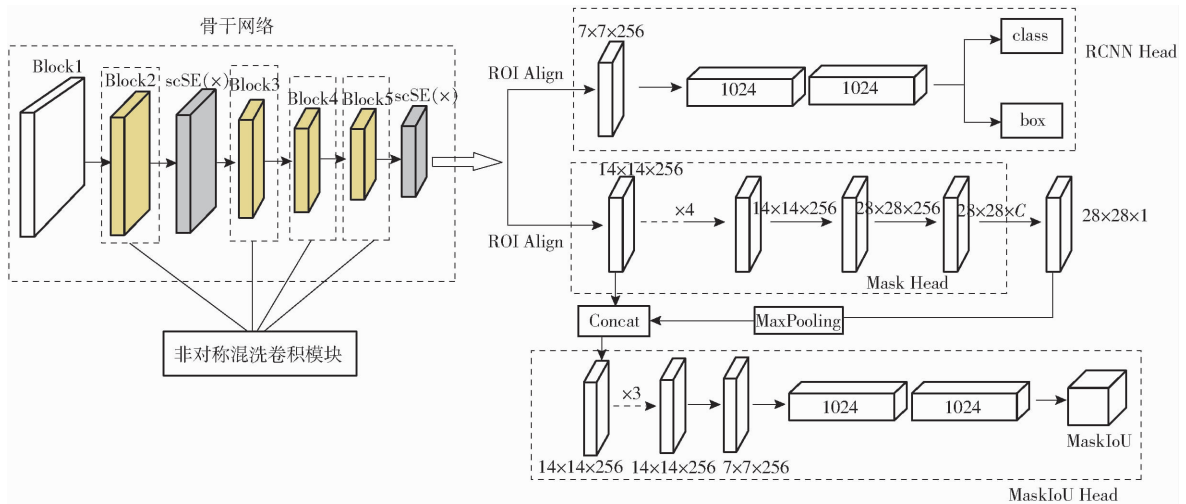


图 1 ASNet 网络结构
Fig. 1 ASNet network structure

2.2 通道压缩和通道混洗

2.2.1 通道压缩

PlantVillage 数据集样本较少且类别不均衡,在使用 ResNeXt - 50 骨干网络时存在大量的参数冗余,这些冗余的参数会使图像特征利用率降低,由于神经网络需要更多的图像数据来进行训练,对数据量较少的类别而言,极易导致过拟合现象的产生,使网络对该类别的测试精度降低。针对此类问题,采

用通道压缩 (Channel squeeze) 的方法减小卷积层输出的通道数,以此减少网络的参数量,降低网络对输入图像数量的依赖性。

在 ResNeXt - 50 网络中,每个 Block 模块的 1×1 卷积层负责减小或增加特征图的通道数, 3×3 卷积层负责提取图像特征,卷积层参数量 $p_{\text{param_conv}}$ 计算公式为

$$p_{\text{param_conv}} = k_w k_h c_i c_o + c_o \tag{1}$$

式中 k_w, k_h ——卷积核尺寸
 c_i, c_o ——卷积层输入、输出通道数
当卷积核尺寸为 1×1 时,公式可简化为

$$p_{\text{param_conv}} = c_i c_o + c_o \tag{2}$$

由式(2)可以看出,模型参数量由输入图像和输出图像的通道数决定,因此本文对每个 Block 模块中第 1 组 1×1 卷积层使用分组卷积,同时对该卷积层的输出通道进行压缩,数量压缩为原来通道数的 $1/4$ 。通道压缩结构如图 2 所示。

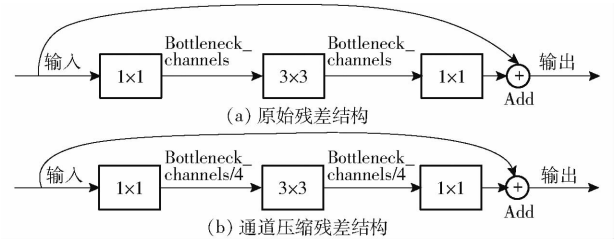


图 2 通道压缩结构示意图

Fig. 2 Structure schematic of channel squeeze

2.2.2 通道混洗

通道混洗(Channel shuffle)^[21]是针对分组卷积阻碍了通道间信息流动的问题所提出的,将通道混洗层放于分组卷积层之后,能有效提高分组卷积中不同通道间的关联性,其结构如图 3 所示。通道混洗模块的操作过程如下:

- (1)将输入的特征图分为 g 组($g_1, g_2, \cdots, g_g$), n 为每组通道数量,输入特征矩阵向量则变换为 (g, n) 。
- (2)进行转置操作变换为 (n, g) 。
- (3)再将得到的结果平坦扁平化,最后结果重新分为 n 组并拼接形成新的特征图作为输出。

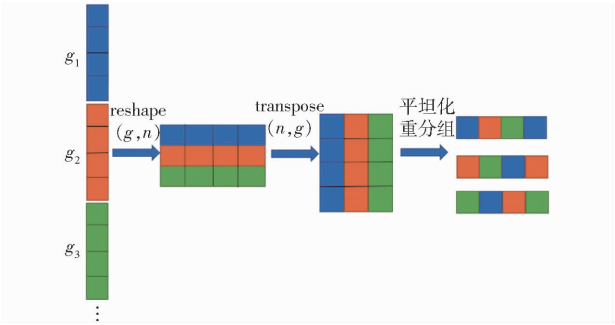


图 3 通道混洗示意图

Fig. 3 Schematic of channel shuffle

通道压缩将 1×1 卷积核的输出通道数降低为原来的 $1/4$,使后续的 3×3 和 1×1 卷积核计算所需要的参数量大幅减少。网络减少了大量的冗余参数,对数据量较少的类别而言,过拟合产生的影响将会极大削弱。同时,分组卷积虽然大幅度降低了参数量,但阻碍了通道间的信息流动,因此使用通道混洗加强各分组卷积之间的联系,使网络以较少的参

数量更好地拟合提取的图像特征,获得更高的模型精度,即使训练集样本总量较少,类别数据量差距较大,依然可以取得较高的分割精度。相较于常规的数据扩充和数据增强而言,极大减少了人工和时间成本。

2.3 非对称混洗卷积模块

在神经网络的使用中,将图像进行翻转或旋转后, $k \times k$ 卷积层提取的图像特征会产生变化,模型对同一目标的分割结果可能会产生偏差,因此导致模型的泛化能力降低,而在卷积神经网络中,模型的特征提取能力和泛化能力是影响最终实验精度的重要因素。

针对上述问题,文献[22]提出了一种非对称卷积结构(Asymmetric convolution net, ACNet),该模块从特征表达角度切入,提升模型精度,在增强模型泛化能力的同时没有带来额外的超参数,推理阶段也没有增加计算量和额外推理时间。作者在训练阶段使用 $k \times k, k \times 1$ 和 $1 \times k$ 3 组并行的卷积层代替 $k \times k$ 卷积层,最后将 3 组卷积层的计算结果进行逐像素相加以得到新的输出特征。在推理阶段,3 组并行卷积层的卷积核进行融合,得到和原始卷积层尺寸相同的 $k \times k$ 卷积核。因此,在推理阶段,网络采用融合后的卷积核,即采用特征提取能力更强的卷积核来提取特征,从而提高精度。

由于推理阶段网络结构与原始网络结构相同,所以没有增加额外计算量。但是在训练阶段,非对称卷积增加的 $k \times 1, 1 \times k$ 卷积层分支,将原始卷积层变为 3 组并行的卷积层,极大增加了网络参数量,使得训练时间大幅度延长,模型时空复杂度也相应的增加。针对此问题,本文在非对称卷积结构 ACNet 的基础上,对该模块进行了优化,提出了非对称混淆卷积。

非对称混洗卷积将非对称结构中的 3×3 卷积层用深度可分离卷积(Depthwise separable convolution, DWConv)^[23]代替,同时,由于图像中的目标尺寸占比较大,且病害的病斑特征区域分布较为分散,因此使用扩张卷积增大感受野以获取更丰富的特征信息。深度可分离卷积分为逐通道卷积和逐点卷积。逐通道卷积是将输入特征图的每个通道独立地进行卷积运算,最后将每个通道的输出特征拼接起来作为结果。这种分组卷积大量减少网络的参数量,但是在相同空间位置上未能充分利用不同通道的特征信息;逐点卷积则是使用 1×1 的卷积层在逐通道卷积的结果上进行深度方向的加权组合,生成新的图像特征作为最终输出结果。

非对称结构中的 1×3 和 3×1 卷积层主要用于

提升模型对翻转或旋转的泛化能力,对空间信息的依赖性降低,因此使用分组卷积能最大限度地减小模型参数。同时,为了加强不同通道间的关联性,对卷积后的结果进行通道混洗。

非对称混洗卷积在非对称卷积的基础上对训练阶段进行了改进,即利用了非对称结构特征提取能力强和泛化能力强的优势,又缩短了模型的训练时

间,减少了网络参数量。非对称混洗卷积模块结构如图 4 所示,图中, dilation 为 Pytorch 卷积操作参数,当 dilation 为 2 时,卷积方式为膨胀卷积。使用该模块替换 Block2、Block3、Block4 和 Block5 中的残差模块。Block2 中的第 1 个非对称混洗卷积模块以 Block1 输出的特征图作为输入,再输出到后续堆叠的相同结构中。

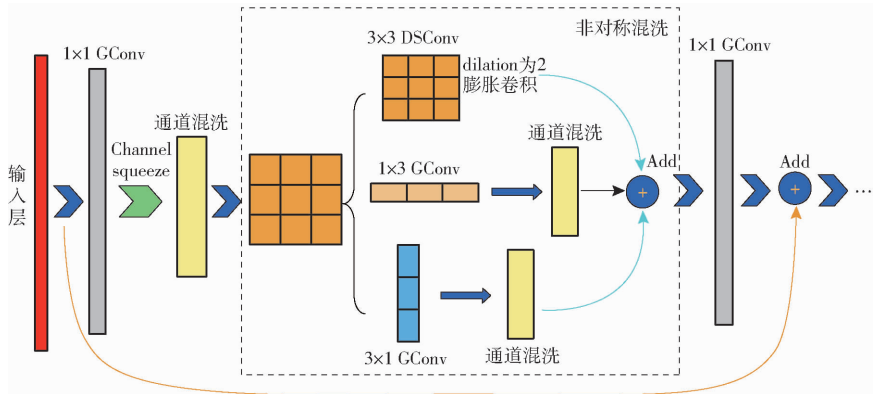


图 4 非对称混洗卷积模块结构图

Fig. 4 Asymmetric shuffle convolution mouldle structure diagram

2.4 注意力机制

卷积神经网络的注意力机制目前已经广泛应用于深度学习的各个领域,成为深度学习技术中重点研究的关键技术之一。注意力机制的本质是从众多的图像信息中选择出对当前任务更关键的信息,并放大这部分信息的权重,抑制其他无用信息的权重,在牺牲有限的计算资源下获得较好的精度。

空间通道压缩与激励注意力机制 (Spatial-channel squeeze & excitation, scSE)是 ROY 等^[24]在压缩与激励注意力机制 (Squeeze-and-excitation, SE)^[25]基础上进行的改进,SE 思想是通过学习输入图像的特征权重并扩大通道中有效特征的权重,减小无效特征权重,以提升模型的精度。scSE 则是在原有的基础上增加了一组空间注意力机制的分支,它通过压缩通道信息以获取空间重要性的指标。

scSE 分为 2 个分支:空间压缩及通道激励分支 (Spatial squeeze and channel excitation block, cSE)和通道压缩及空间激励分支 (Channel squeeze and spatial excitation block, sSE)。

cSE 是通道注意力机制,该分支首先将输入特征图进行全局池化,然后通过一组全连接层将输入特征图变换为 $1 \times 1 \times C/2$ 的张量,其中 C 为输入特征图通道数,再经过激活函数 (ReLU),最后通过一组全连接层变换为 $1 \times 1 \times C$ 的特征向量,再由归一化层 (Sigma),将每个值归一化到 (0,1) 之间。cSE 的实际意义是将每个通道的全部信息压缩为一个 0 到 1 之间的值,再将这些值分别乘到这些信息对应

的通道上,增大重要特征信息的通道权重。

sSE 是空间注意力机制,主要的操作是通过一组 1×1 的卷积层来实现空间压缩操作,然后通过 Sigma 层实现归一化,再将这些值乘以对应的输入矩阵,扩大或缩小不同空间信息的权重。sSE 的实际意义是确定重要特征信息的空间位置并赋予更大的权重。

scSE 将 cSE 与 sSE 的输出特征进行逐像素相加,由于两组分支的输出结果均为归一化后的特征矩阵,因此在增加有用信息数值的同时,在一定程度上也会增大无用信息的数值,2 种信息的权重比值在改变前后的差距上所拉大的程度并不多。因此,本文通过将分支结果进行逐像素相乘,分支间的逐像素相乘可以让无用信息的权重相较于有用信息的权重占比更小,更利于排除无用的信息,经过实验,相较于原始 scSE,改进后的注意力机制精度有小幅提升。

3 实验结果及分析

3.1 图像采集和数据集建立

实验数据集分为两部分,一部分来自公开的 PlantVillage 数据集,另一部分来自 2018 年 AI Challenger 农作物病害分割挑战赛和实地采集的苹果叶部病害图像。

对于公开的 PlantVillage 数据集,本文选用了苹果叶片 RGB 图像 1 325 幅,包括黑腐病 358 幅、锈病 275 幅、黑星病 404 幅和健康叶片 288 幅。将以上

图像按照 7:3 的比例划分为训练集和测试集。

将农作物病害分割挑战赛和实地采集的苹果叶部病害图像作为混合数据集,共 240 幅图像,包括每类单目标图像 60 幅,用作第 2 测试集以确定模型的泛化能力。

其中实地采集的图像 103 幅,采集的环境包括不同光照条件下的病害图像,包括强光、正常光照和阴天。实地采集图像使用荣耀 V30pro 型手机内置相机,像素为 4 000 万,采集图像尺寸为 2 592 像素 × 4 608 像素,在采集过程中,应保证病害叶片的病害区域完全包括在图像中,并尽量扩大该病害叶片在整幅图像中的占比,尽可能排除背景颜色和不同叶片相互遮挡的干扰。

农作物病害分割挑战赛图像 137 幅,由于挑战赛的图像与 PlantVillage 数据集有一定的相似性,因此将挑战赛获取的图像进行旋转和翻转,用于测试本文模型的鲁棒性。

PlantVillage 数据集和混合数据集如图 5 所示,其中图 5a ~ 5d 为 PlantVillage 数据集获取的图像,图 5e ~ 5g 为混合数据集图像。

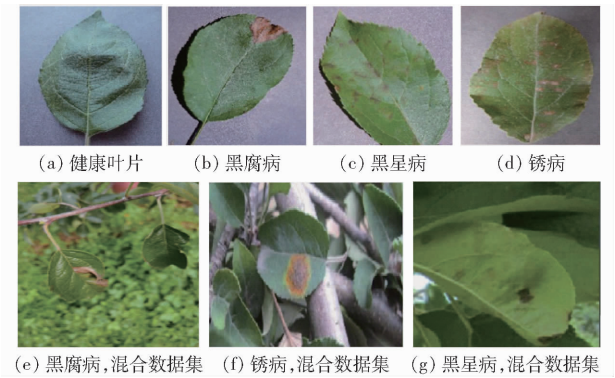


图 5 PlantVillage 数据集和混合数据集苹果叶片病害示例
Fig. 5 Examples of apple leaf diseases in PlantVillage dataset and mixed dataset

3.2 实验环境和超参数

实验环境为 Ubuntu16.04 操作系统,Python 3.6,深度学习框架为 Pytorch 1.0.0。硬件环境 CPU 为 AMD R5 - 3600,16 GB 内存, GPU 为 RTX2060 6GB 显卡,使用 CUDA Toolkit 9.0 和 CUDNN V7.0.5 作为网络模型训练加速工具包。

为确定最佳的训练参数,在实验中分别采用不同的学习率对 ASNet 模型进行训练,同时输出不同学习率下的损失函数进行对比,实验结果如图 6 所示。对实验结果分析可得,学习率为 0.01 时,收敛的速度最快,并且收敛的结果最好。学习率为 0.005 时,收敛的速度变慢,且在训练过程中波动起伏较大。而学习率继续减小时,最终收敛的数值较大,网络陷入局部最优。

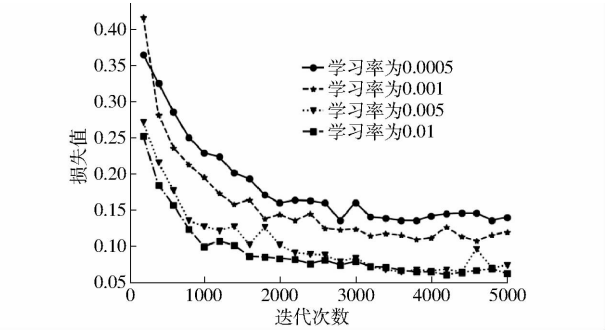


图 6 不同学习率下损失曲线对比
Fig. 6 Comparison of loss curves at different learning rates

输入批尺寸 (Batchsize) 应当在显存足够的情况下尽量增大。增大输入批尺寸可以加快网络对数据量的处理速度,并且在训练过程中,损失函数的振荡更小。因此, Batchsize 设置为 4。各超参数如表 1 所示。

表 1 超参数	
Tab.1 Determination of hyperparameters	
超参数	策略及数值
初始学习率	0.01
Weight_decay	0.000 1
学习率衰减策略	Step
Decay_steps	(4 000,7 000)
Gamma	0.1
热身迭代次数	500
热身迭代策略	Liner
迭代总数	10 000

3.3 评价指标

COCO 评价指标 (Evaluation metric) 与 VOC 系列数据集使用的评价指标对平均精度 (Average precision, AP) 要求的阈值不同。VOC 数据集使用的阈值,即交并比 (Intersection over union, IoU) 满足 0.5 即可。而 COCO 数据集使用的阈值指标以 0.05 为增量,分别计算 0.5 ~ 0.95 共 10 个阈值下的 AP,最后将所有阈值的 AP 求和取平均。COCO 评价指标对实验结果的准确性要求远高于 VOC 评价指标,并且更能体现算法的真实性能。

基于 Mask R - CNN 改进,实验结果采用 COCO 评价指标:平均精度 (Mean average precision, mAP) 和平均召回率 (Mean average recall, mAR)。

3.4 实验

3.4.1 骨干网络对比实验

为了确定本文模型骨干网络的选择,在 Mask R - CNN 的框架下,分别对骨干网络为 ResNet - 50、ResNet - 101、ResNeXt - 50 和 ResNeXt - 101 的模型进行了实验对比,实验结果如表 2 所示。

ResNeXt 相较于 ResNet 增加了网络的宽度,精

表 2 不同骨干网络实验结果

Tab. 2 Experimental results of different backbone networks

骨干网络	平均精度/ 模型权重文		参数量	测试 时间/s
	%	件大小/MB		
ResNet-50	91.3	334	4.30×10^7	0.045
ResNeXt-50	91.6	491	6.10×10^7	0.055
ResNet-101	92.2	501	6.20×10^7	0.057
ResNeXt-101	84.9	815	1.02×10^8	0.070

度有小幅提升;ResNet-101 通过加深网络层深度使得网络的模型权重文件在增加了 167 MB 后精度提升了 0.9 个百分点。

ResNeXt-50 和 ResNet-101 在测试精度和模型最终权重文件相差较小。但是 ResNeXt 在结构上的改进相对于 ResNet 具有更强的表征能力,并且这种分组操作也能起到一定正则化的作用,使卷积层学习到的特征更加稀疏。因此,本文后续所有实验骨干网络均采用 ResNeXt-50。

3.4.2 改进注意力机制实验

为了验证注意力机制模块数量及其使用位置对苹果叶片病害数据集的影响,模型采用 ResNeXt-50 骨干网络,在 Mask R-CNN 的各 Block 模块之后添加本文改进的注意力机制模块。实验结果如表 3 所示。

表 3 不同位置添加注意力机制的实验结果

Tab. 3 Experimental results of adding attention mechanism to different positions

位置	平均精度	平均召回率
Block2 + 3	91.6	95.3
Block2 + 4	92.1	96.7
Block2 + 5	92.7	96.5
Block3 + 4	92.4	96.9
Block3 + 5	90.7	95.8
Block4 + 5	92.1	96.3
All Blocks	91.6	96.3

由表 3 可以看出,对于本文的数据集而言,注意力机制的个数不宜添加过多,最终将注意力机制添加在具有残差结构的第 1 个模块和最后 1 个模块之后,即 Block2 和 Block5,实验取得的结果最好,精度最高达到 92.7%,相比于原模型精度 91.3%,使用改进的注意力机制模块后,模型精度提升了 1.4 个百分点。

分析原因为:Block2 作为第 1 个具有残差结构的模块,在它输出的特征图上进行权重的扩大和抑制,对后续的 Block 都会产生有利的影响。Block5 输出的特征图具有最抽象的语义,对它进行权重的重新分配则对最终的预测结果会产生积极的影响。

因此,只添加 2 个注意力机制模块能在增加有限的计算量的情况下获得更好的结果。

3.4.3 消融实验

为了验证各改进结构在本文小样本数据集中对病害分割的有效性,在 Mask R-CNN 的框架下,将本文模型 ASNet 分别与添加了 MaskIoU Head (Mask Scoring R-CNN, MS R-CNN)、改进注意力机制 scSE(×) 和非对称混洗卷积模块的不同网络进行实验,实验结果如表 4 所示。

表 4 不同结构下苹果叶部病害分割结果

Tab. 4 Segmentation results of apple leaf diseases by different networks

模型	MaskIoU Head	scSE(×)	非对称混洗模块	mAP/%
Baseline				91.6
MS R-CNN	■			91.9
本文模型		■		92.7
本文模型			■	93.5
本文模型	■	■		94.3
本文模型		■	■	94.1
本文模型	■		■	95.5
本文模型	■	■	■	96.8

注:■表示在网络中添加相应结构。

由表 4 可知,MaskIoU Head 作为新加分支,虽然修正了 Mask 与 Mask 得分之间的偏差,但带来了大量额外的冗余参数,因此最终精度仅提升了 0.3 个百分点;改进后的注意力机制在增大了对无用特征信息的抑制力度后,精度提升了 1.1 百分点;在单独添加非对称混洗卷积模块时,3 组并行分支和扩大感受野的结构能有效提高病害分割精度,并且精度提高最大,达到 1.9 个百分点。

本文模型在增加了 MaskIoU Head 分支、改进注意力机制和非对称混洗卷积结构后,精度提升至 96.8%,相较于 Mask R-CNN 提升了 5.2 个百分点。表明本文模型的特征提取能力和对图像旋转、翻转的抗干扰能力均有所增强,从而使模型提升的精度最多。

3.4.4 通道混洗实验

为验证通道混洗添加的数量及其位置对苹果叶片病害分割的影响,模型采用 ResNeXt-50 作为骨干网络,在 ASNet 的框架下对通道混洗添加的数量及位置进行比较分析。

通道混洗能有效提高苹果叶部病害的平均分割精度,并且在不同位置添加该模块对实验精度影响较大,通过实验可得,在所有使用分组卷积的卷积层后使用通道混洗对测试精度的提升最大,可提升 2.2 个百分点。通道混洗加强了通道间特征信息的

流动,使模型在分组卷积后对不同通道中处于相同空间位置的特征信息利用效率提高,从而提高实验精度。

3.4.5 扩张卷积实验

为了验证扩张卷积的扩张率对苹果叶片病害分割精度的影响,模型采用 ResNeXt-50 骨干网络,在 ASNet 框架下,对非对称混洗卷积模块的 3×3 卷积层扩张率进行比较分析。

非对称混洗卷积模块中使用了扩张卷积来使卷积核获得更大的感受野。当扩张率为 2 时,模型取得的精度最高达到 96.8%;当扩张率为 4 时,测试精度为 95.5%,仍然高于原始网络基准 95.2%;但继续增大扩张率时,精度开始逐渐下降。实验结果表明当扩张率大于 4 时,卷积核的感受野虽然扩大,但信息丢失较为严重,导致实验精度不断降低。

3.4.6 不同网络模型实验对比

PlantVillage 数据集中的苹果叶部病害图像存在数据样本少、目标类别数不均衡的问题;混合数据集中的图像则包含了光照强度不同、拍摄角度不同的情况。为了验证本文提出的卷积神经网络模型 ASNet 在小样本数据集下对苹果叶部病害分割的优越性,分别使用目前较为成熟和先进的 Mask R-CNN、Yolact^[26]、Yolact++^[27]和 SipMask^[28]在 2 个数据集上进行实验,实验结果如表 5 所示。

表 5 不同网络实验结果对比

Tab.5 Comparison of experimental results for different networks %

任务网络	PlantVillage 数据集		混合数据集	
	平均精度	召回率	平均精度	召回率
Mask R-CNN	91.6	95.9	87.0	93.8
Yolact	95.1	96.6	82.5	90.3
Yolact++	95.6	96.7	83.4	91.4
SipMask	87.7	98.7	74.6	93.8
ASNet	96.8	98.4	94.7	94.9

Yolact 和 Yolact++ 在图像背景相对简单的 PlantVillage 数据集上,均取得了较高的精度,并且网络实时性达到 30 帧/s 以上,但是在混合数据集上的测试精度出现大幅度降低,表明模型的泛化能力相对较弱;SipMask^[28]是最新的基于空间信息保存的单阶段实例分割算法,该网络在 COCO 数据集上精度和实时性均超过了 Mask R-CNN,但面对本文的小样本数据集时,在 2 个数据集上的测试精度都相对较低。

本文提出的 ASNet 在样本少、类间数量不均衡的 PlantVillage 数据集和成像条件具有差异的混合

数据集上均取得了最好的结果,平均测试精度分别达到 96.8% 和 94.7%。因此本文模型的改进能有效降低网络对输入图像数量的依赖性,降低类间数量不均衡对实验精度的影响,增强模型泛化能力。

3.4.7 混合数据集测试实验

进一步验证本文模型能有效降低过拟合产生的影响,减少模型对输入图像数量的依赖性。骨干网络均采用 ResNeXt-50,分别使用 Mask R-CNN、MS R-CNN 和 ASNet 对混合数据集的每类病害进行测试,实验结果如表 6 所示。

表 6 混合数据集的类间分割精度

Tab.6 Inter-class segmentation accuracy of mixed data sets

任务网络	分割精度/%					模型权重文件大小/MB
	黑腐病	健康	锈病	黑星病	平均	
Mask R-CNN	83.1	90.5	82.0	92.3	87.0	491
MS R-CNN	84.4	94.3	77.2	92.8	87.2	616
ASNet	91.1	98.7	93.8	95.2	94.7	321

由表 6 可知,对混合数据集平均精度而言,在使用增加了 MaskIoU Head 分支的 MS R-CNN 进行测试时,平均精度并未产生较大的变化,说明网络中存在大量的冗余参数;对每种病害类别而言,由于 PlantVillage 数据集中锈病(rust)数据量最少,使网络对该类别的训练学习效果最差,所以其分割精度也最低;与 Mask R-CNN 实验结果相比,在所有类别精度均上升的情况下,锈病的分割精度却出现较大幅度的下降,说明对数据较少的类别而言,这些冗余参数会导致过拟合现象造成的影响加大,从而使病害分割的平均精度降低。

在使用本文提出的卷积神经网络模型 ASNet 进行测试时,所有类别的分割精度均有大幅度提升,其中锈病提升最大,并且该类的精度也不再是最低的类别。实验表明:在本文模型中,实验精度更加依赖于网络对类别特征提取的准确性,降低了类别数据量对实验结果的影响,削弱了由于样本量少导致的过拟合现象对精度的影响。

由 ASNet 的实验结果可知,其对健康叶片的分割精度最高,达到 98.7%,但对黑腐病和锈病的分割精度较低,分别为 91.1% 和 93.8%,其中黑腐病不同时期的叶片病害特征差距较大,早期特征为紫色圆形病斑,广泛分布在叶片的不同区域,后期病斑边缘隆起,中部凹陷,边缘呈为黑褐色,中部呈暗灰色,并且多个病斑密集连形成不规则大块病斑,从而导致黑腐病的特征提取难度增大;锈病初期表面会产生小点状性孢子器,为橙黄色,后期病斑周围产生

毛状锈孢子器,颜色也变为黄褐色,但锈病后期出现的子器及孢子,特征较为明显,因此分割精度高于黑腐病。黑星病发病初期叶片上会呈现出绿褐色霉层,随后霉层渐变为褐色至黑色,由于不同时期的黑星病均存在覆霉层,因此对该病害的分割精度最高,达到 95.2%。

模型最终分割结果如图 7 所示,可以看出,改进算法对苹果叶部病害分类准确,同时定位效果较好,掩码质量较高。

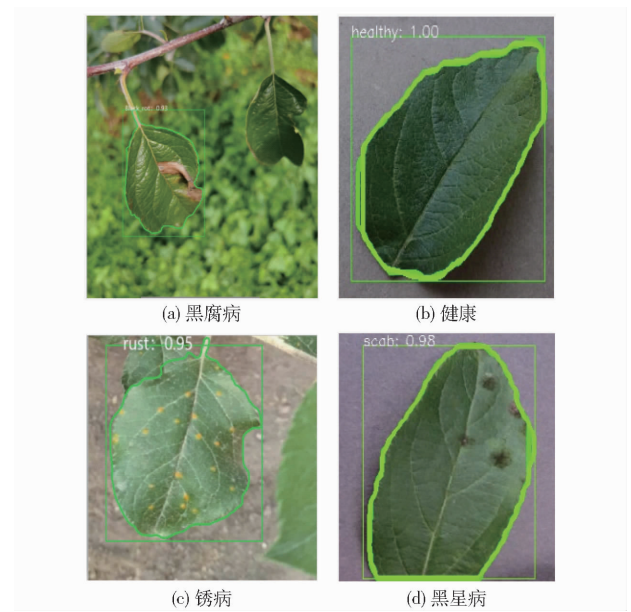


图 7 分割结果可视化示例

Fig. 7 Visualization example of test results

4 结论

(1)提出了一种非对称混洗卷积网络 ASNet 以作为苹果叶部病害高精度分割的模型,为深度学习在苹果叶部病害分割上的应用提供了一定参考价值。

(2)网络在 Mask R - CNN 基本框架下,对骨干网络 ResNeXt - 50 进行了改进。用非对称混洗卷积代替传统 3×3 卷积核,同时针对小样本的数据集采用了通道压缩和通道混洗的方法减小网络对输入图像数据量的依赖性,并利用改进 scSE 注意力机制增强网络提取的特征,在增加精度的同时减小了模型的时空复杂度。

(3)提出的卷积神经网络模型 ASNet 对 PlantVillage 数据集图像的分割精度最高达到 96.8%,对混合的数据集图像的分割精度最高达到 94.7%。模型在 2 种数据集上均取得了较好的结果,因此模型的泛化能力较强。

(4)本研究为数据集样本总量少、类间样本不均衡导致的常规深度学习模型精度偏低的问题提供了参考方案,在兼顾模型精度和轻量化两方面选择采用不同的卷积方法,即提出的非对称混洗卷积模块,通过扩张卷积扩大了卷积核感受野,提升了模型对病斑特征较为分散的病害叶片的特征提取能力;同时使用并行非对称结构增强了网络对图像旋转和翻转的鲁棒性,这为苹果叶部病害的分割提供了一定参考价值。

参 考 文 献

[1] 王树桐,王亚南,曹克强. 近年我国重要苹果病害发生概况及研究进展[J]. 植物保护,2018,44(5):13 - 25.
WANG Shutong, WANG Ya'nan, CAO Keqiang. Occurrence of and research progress in important apple diseases in China in recent years[J]. Plant Protecti, 2018, 44(5): 13 - 25. (in Chinese)

[2] 牟悦龙. 浅谈苹果树的病害防治对策和措施[J]. 农业与技术,2018,38(20):35.

[3] 索相敏,冯少菲,郝敏. 几种常见苹果叶部病害及其防治技术[J]. 河北果树,2018(6):24 - 25.

[4] 王建玺,宁菲菲,鲁书喜. 基于支持向量机的苹果叶部病害识别方法研究[J]. 山东农业科学,2015,47(7):122 - 125.
WANG Jianxi, NING Feifei, LU Shuxi. Study on apple leaf disease recognition method based on support vector machine[J]. Shandong Agricultural Sciences, 2015, 47(7): 122 - 125. (in Chinese)

[5] 石凤梅,赵开才,孟庆林,等. 基于支持向量机的水稻稻瘟病图像分割研究[J]. 东北农业大学学报,2013,44(2):128 - 135.
SHI Fengmei, ZHAO Kaicai, MENG Qinglin, et al. Image segmentation of rice blast in rice based on support vector machine [J]. Journal of Northeast Agricultural University, 2013, 44(2): 128 - 135. (in Chinese)

[6] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术研究综述[J/OL]. 农业机械学报,2021,52(7):1 - 18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1 - 18. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20210701&flag=1. DOI:10.6041/j.issn.1000-1298.2021.07.001. (in Chinese)

[7] 许景辉,邵明烨,王一琛,等. 基于迁移学习的卷积神经网络玉米病害图像识别[J/OL]. 农业机械学报,2020,51(2):230 - 236, 253.
XU Jinghui, SHAO Mingye, WANG Yichen, et al. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230 - 236, 253. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=2002025&flag=1. DOI: 10.6041/j.issn.1000-1298.2020.02.025. (in Chinese)

[8] 张善文,王振,王祖良. 多尺度融合卷积神经网络的黄瓜病害叶片图像分割方法[J]. 农业工程学报,2020,36(16):149 - 157.
ZHANG Shanwen, WANG Zhen, WANG Zuliang. Method for image segmentation of cucumber disease leaves based on multi-

- scale fusion convolutional neural networks[J]. Transactions of the CSAE, 2020, 36(16):149–157. (in Chinese)
- [9] 何欣,李书琴,刘斌. 基于多尺度残差神经网络的葡萄叶片病害识别[J]. 计算机工程,2020,47(5):8.
HE Xin,LI Shuqin,LIU Bin. Grape leaf disease identification based on multi-scale residual network[J]. Computer Engineering, 2020,47(5):8. (in Chinese)
- [10] 郭小清,范涛杰,舒欣. 基于改进 Muti-scale AlexNet 的番茄叶部病害图像识别[J]. 农业工程学报,2019,35(13):162–169.
GUO Xiaoqing, FAN Taojie, SHU Xin. Tomato leaf diseases recognition based on improved Multi-scale AlexNet[J]. Transactions of the CSAE, 2019, 35(13):162–169. (in Chinese)
- [11] 张乃夫,谭峰,范禹希,等. 基于卷积神经网络的农作物病害识别方法研究[J]. 安徽农业科学,2020,48(5):242–245.
ZHANG Naifu,TAN Feng,FAN Yuxi, et al. Research on crop disease identification method based on convolutional neural network[J]. Journal of Anhui Agricultural Sciences,2020,48(5):242–245. (in Chinese)
- [12] 邱靖,刘继荣,曹志勇,等. 基于卷积神经网络的水稻病害图像识别研究[J]. 云南农业大学学报(自然科学版),2019, 34(5):884–888.
QIU Jing,LIU Jirong,CAO Zhiyong, et al. Rice disease image recognition research based on convolutional neural network[J]. Journal of Yunnan Agricultural University (Natural Science),2019,34(5):884–888. (in Chinese)
- [13] 黄建平,陈镜旭,李克新,等. 基于神经结构搜索的多种植物叶片病害识别[J]. 农业工程学报,2020,36(16):166–173.
HUANG Jianping, CHEN Jingxu, LI Kexin, et al. Identification of multiple plant leaf diseases using neural architecture search [J]. Transactions of the CSAE, 2020, 36(16):166–173. (in Chinese)
- [14] 吴娜. 基于图像检索的农作物叶部病害识别关键技术研究[D]. 合肥:中国科学技术大学,2018.
WU Na. Research on key techniques of crop leaf diseases recognition based on image retrieval[D]. Hefei:University of Science and Technology of China,2018. (in Chinese)
- [15] XIER S, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:1492–1500.
- [16] SZEGED Y, LIU W, JIA Y, et al. Going deeper with convolutions[C] //Proceedings of International Conference on Computer Vision and Pattern Recognition. IEEE, 2015: 1–9.
- [17] HE Kaiming, ZHANG Xianyu, REN Shaoqing, et al. Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[C] //Neural Information Processing Systems (NIPS), Lake Tahoe, 2012: 1097–1105.
- [19] HE Kaiming, GKIOXARI G, DOLLAR P, et al. Mask R – CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 42(99):386–397.
- [20] HUANG Zhaojin,HUANG Lichao,GONG Yongchao,et al. Mask Scoring R – CNN [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 6409–6418.
- [21] ZHANG Xiangyu, ZHOU Xinyu,LIN Mengxiao, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848–6856.
- [22] DING Xiaohan, GUO Yuchen, DING Guiguang, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C] // Proceedings of the IEEE International Conference on Computer Vision, 2019:1911–1920.
- [23] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. Proceedings of the IEEE, 2017,105(12): 2295–2329.
- [24] ROY A G, NAVAB N, WACHINGER C, et al. Concurrent spatial and channel ‘Squeeze & Excitation’ in fully convolutional networks[C] // Proceedings of International Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 6848–6856.
- [25] HU J, LI S, GANG S, et al. Squeeze-and-excitation networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132–7141.
- [26] BOLYA D, ZHOU C, XIAO F, et al. YOLACT: real-time instance segmentation[C] // Proceedings of IEEE International Conference on Computer Vision, 2019: 9157–9166.
- [27] BOLYA D, ZHOU C, XIAO F, et al. YOLACT + + : better real-time instance segmentation[J]. arXiv preprint arXiv:1912.06218,2019.
- [28] CAO Jiale, ANWER R M, CHOLAKKAL H, et al. SipMask: spatial information preservation for fast image and video instance segmentation[C] //Proceedings of the European Conference on Computer Vision, 2020.