

DOI:10.16644/j.cnki.cn33-1094/tp.2020.10.018

基于Mask R-CNN的宫颈细胞图像分割

郑 杨¹, 梁光明², 刘任任¹

(1. 湘潭大学计算机学院·网络空间安全学院, 湖南 湘潭 411105; 2. 国防科技大学电子科学与工程学院)

摘要: 宫颈细胞图像中目标分割的精度直接影响对疾病的判别和诊断, 宫颈细胞图像中有不同种类的多个目标, 所以有必要对宫颈细胞图像进行实例分割。为了获得更好的宫颈细胞图像实例分割效果, 文章在Mask R-CNN的基础上提出了一种宫颈细胞图像实例分割方法, 在网络中的特征金字塔网络(FPN)中加入空洞卷积将其改造为DFPN, 减少图像信息的损失来提升分割的准确度。在TCTCOCO数据集的测试结果表明, 该方法提高了宫颈细胞图像分割的精度。

关键词: 卷积神经网络; 宫颈细胞显微图像; 实例分割; 空洞卷积

中图分类号: TP391

文献标识码: A

文章编号: 1006-8228(2020)10-68-05

Cervical cell image segmentation based on Mask R-CNN

Zheng Yang¹, Liang Guangming², Liu Renren¹

(1. School of Computer Science, Xiangtan University, Xiangtan, Hunan 411100, China; 2. School of Electronic Science and Engineering, National University of Defense Technology)

Abstract: The accuracy of object segmentation in cervical cell image directly affects the identification and diagnosis of the disease. There are many different kinds of objects in cervical cell image, so it is necessary to segment cervical cell image by instance. In order to obtain a better instance segmentation effect of cervical cell image, this paper proposes an instance segmentation method of cervical cell image based on Mask R-CNN. The feature pyramid network (FPN) in the network is added with dilated convolution to transform it into DFPN to reduce the loss of image information and improve the accuracy of segmentation. The results of test on TCTCOCO data set show that the method improves the accuracy of cervical cell image segmentation.

Key words: convolutional neural network (CNN); cervical cell microscopy image; instance segmentation; dilated convolution

0 引言

许多疾病的确诊是依靠医生分析医学影像, 但是医生的经验和劳累程度会影响诊断结果, 因此, 需要利用计算机视觉相关技术来辅助医生得出诊断结果, 这样可以提高效率 and 降低误诊率。

随着智能医疗的发展和相关政策出台, 已有越来越多的人关注和研究智能医疗。如今医学图像分割的方法分为两类, 一类是传统方法, 其主要包括基于边缘的方法^[1]、基于阈值的方法^[2]、基于聚类的方法^[3]、基于区域的方法^[4]; 另一类是基于卷积神经网络的方法^[5], 其通过卷积进行特征提取, 然后再进行分割。以上方法仅对图像进行语义分割, 而实例分割是目标检测和语义分割的结合, 既能分割得到物体边缘, 又能标出图像中相同种类物体中不同的个体。目标检测

的方法现今主要分为一步法和二步法两类: Yolo^[6]和SSD^[7]都是一步法; 二步法有Fast R-CNN^[8]、Faster R-CNN^[9]和Mask R-CNN^[10]。二步法对于一步法而言, 虽然速度慢, 但是精确度较高。

宫颈细胞图像的处理结果直接影响诊断, 所以本文选取二步法中目标检测精度最高且能进行分割的Mask R-CNN。宫颈细胞图像中的极小目标在特征提取时, 可能产生信息损失, 且损失无法避免也不可逆, 减少图像信息的损失是提高分割准确度来说很关键。为了得到更好的结果, 本文也通过这一思路来对Mask R-CNN进行改进。

1 方法

1.1 Mask R-CNN简介

Mask R-CNN是在Faster R-CNN上添加了分割

收稿日期: 2020-05-06

作者简介: 郑杨(1994-), 男, 湖南湘西人, 硕士, 主要研究方向: 图像处理。

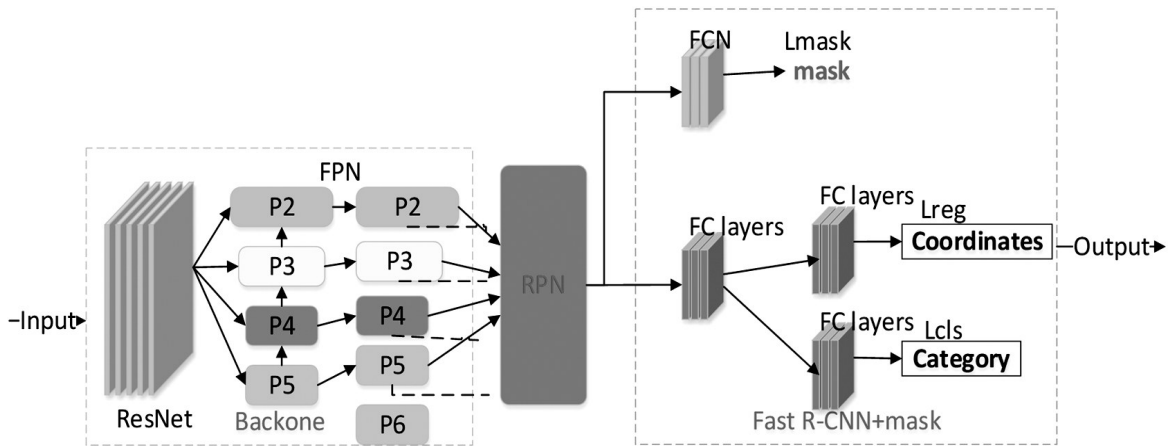


图1 Mask R-CNN 结构图

的分支,对目标候选框区域(ROI)中的目标使用全卷积网络(FCN)^[11]进行分割。其结构图如图1所示。

Mask R-CNN的损失函数分为三个部分:

(1) RPN^[9]: RPN对ROI做前景和背景分类损失和回归损失,损失函数为:

$$L_{RPN} = L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

其中, N_{cls} (每个batch中锚框的数量) 和 N_{reg} (锚框位置的数量) 是对两项进行归一化, 平衡参数 λ 是对其进行加权。在一个batch中的锚框的编号为 i , 锚框 i 对象的预测概率为 p_i 。如果锚框预测正确, 则真实标签 (Ground Truth label) p_i^* 为1, 反之则为0。 t_i 代表预测边界框的4个参数化坐标的向量, t_i^* 则对应的是锚框实际边界框的参数化坐标。分类损失 L_{cls} 是两个类别的对数损失(前景和背景)。对于回归损失采用 $smooth_{l_1}$ 损失如公式(2):

$$smooth_{l_1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (2)$$

(2) Fast R-CNN分支: 跟RPN部分类似也是分类和回归损失, 不同的是Fast R-CNN中的分类是对ROI中的目标进行分类的损失。Fast R-CNN部分的损失函数为:

$$L_{Fast R-CNN} = L_{cls} + L_{reg} \quad (3)$$

(3) Mask分支: Fast R-CNN的结果输入Mask分支, 进行ROIAlign操作, 再对每个ROI用FCN对逐个像素点进行softmax分类输出mask, 这一分支的损失即训练生成mask的损失即为 L_{mask} 。总的损失函数为:

$$L_{total} = L_{RPN} + L_{Fast R-CNN} + L_{mask} \quad (4)$$

Mask R-CNN进行实例分割的具体流程如图2。

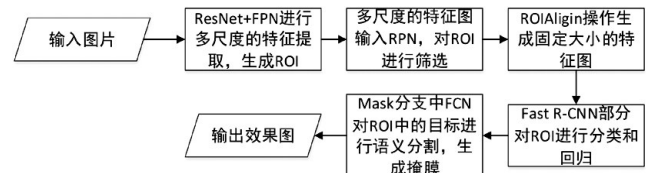


图2 Mask R-CNN 实例分割流程图

1.2 FPN的改进DFPN

FPN^[12]的结构如图3所示, Image通过ResNet^[13]得到P1到P3不同尺度的feature map, 右侧P4到P6的过程中每个阶段都一样。以P5到P6为例, P1经过 1×1 的卷积得到的结果与P5上采样得到的feature map做相加, 最终得到P6, 然后将P6进行 3×3 的卷积后送入RPN, 其他不同尺度层数的操作都是如此。

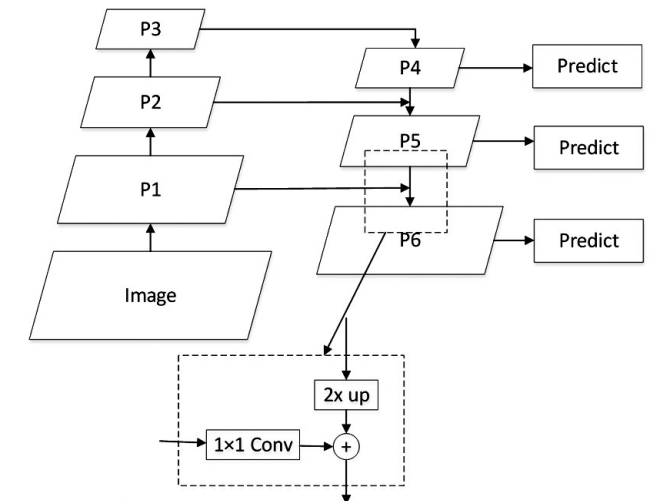


图3 FPN 结构图

图3中P4到P5和P5到P6阶段, 进行上采样会造成图像信息的损失, 而且是不可逆的, 虽然FPN中各层特征图均进行了特征融合来减少图像信息损失, 但是一些小特征容易被忽视, 而空洞卷积^[14]可以通过控

制卷积核rate的数值得到对应不同大小的感受域,图4中均为卷积核是 3×3 的空洞卷积,rate=1时感受域为 3×3 ,rate=2时的感受域变为 7×7 (红色部分为卷积核)。通过不同rate的空洞卷积能获得不同感受域的图像信息,因此本文选择在FPN中加入空洞卷积,让高级层特征与低层特征也进行融合来减少图像信息的损失。

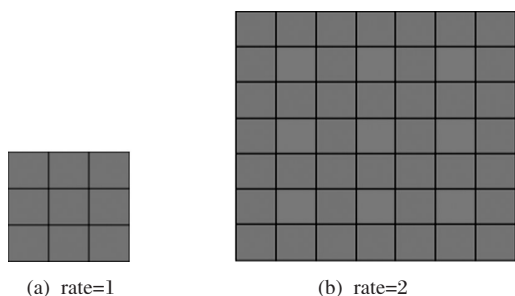


图4 空洞卷积

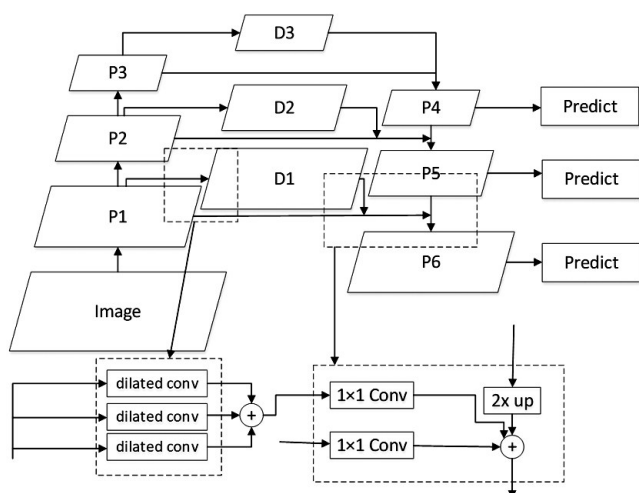


图5 DFPN结构图

空洞卷积在增大了感受野的同时不会减小图像大小,可以捕获更多信息,解决了卷积神经网络(CNN)对保持特征图尺寸的同时增加感受域的问题^[15]。把空洞卷积应用进FPN中,使其在特征融合中可以融合更多的图像信息,从而提高分割的准确度。本文称加入空洞卷积后的FPN为DFPN,结构图如图5所示,以P5到P6为例,P1分别经过rate=2,rate=3,rate=4的卷积核均为 3×3 的三种空洞卷积得到的三种feature map相加得到D1(D1相当于另一种意义上的不同尺度的特征图的融合,但是拥有高级层特征的同时,并没有改变尺寸大小),D1和P1分别都再通过 1×1 卷积的结果与P5上采样的结果进行特征融合得到P6,后续步骤保持不变。

本文方法即使用DFPN代替Mask R-CNN中的

FPN对宫颈显微图像进行实例分割。

2 实验

2.1 TCTCOCO数据集的制作

本文方法训练所需的数据集为COCO数据集^[16],本文基于长沙市第二人民医院提供的宫颈TCT细胞涂片制作了TCTCOCO数据集。具体流程如图6所示。

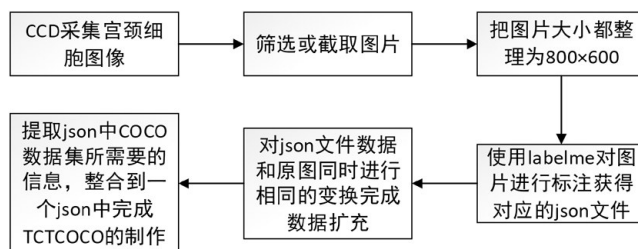


图6 TCTCOCO制作流程图

第二步和第三步是为了方便训练,数据标注在医生和检验人员指导下使用labelme进行,标注图片的数量为2000张,图像中标注的目标为三类,其中上皮细胞标签为“shangpi”;白细胞标签为“bai”;真菌标签为“zhenjun”。制作完成的数据集相关信息存储于后缀名为json的文件中。

为了避免因数据集小产生过拟合,通过人工方法扩充数据集到20000张,训练集、测试集和验证集的分配见表1。

表1 训练集、测试集、验证集数据比例分布

训练集	测试集	验证集	总数
14000(70%)	2000(10%)	4000(20%)	20000

2.2 实验环境

本文所有实验的环境都如下:

CPU: Intel(R) Core(TM) i7-8700 内存: 16GB
显卡: GeForce RTX 2070 SSD: 256GB
硬盘: 4TB 系统: Windows 10

软件环境和库: Anaconda 3-python3.7, cuda9.0, mxnet, Numpy, Opencv-python等。

2.3 实验结果分析

为了保证实验的一致性,本文所有实验均将整个数据集迭代1000次,学习率为0.01,使用GPU进行训练。评价指标为IoU (intersection over union) 和 mIoU (Mean intersection over union),其中,IoU和mIoU定义如下:

$$\text{IoU}(i) = a_{ii} / (\sum_j (a_{ij} + a_{ji}) - a_{ii}) \quad (5)$$

$$\text{mIoU} = \frac{1}{N} \sum_i \text{IoU}(i) \quad (6)$$

其中, a_{ii} 为正确分割的像素点的数目(即类别为 i 的像素点数目被预测为 i 的数目); a_{ij} 为被预测为 j 的 i 类像素点数目; a_{ji} 为 j 被预测为 i 的 j 类像素点数目; N 表示总的类别数。图像部分原图如图 7 中(a)(b)(c)所示, 对应的 Mask R-CNN 效果图如(d)(e)(f)所示, (g)(h)(i)是本文方法生成的效果图。

Mask R-CNN 与本文方法都生成了预测的目标框和分类结果, 每个目标都生成了各色的掩膜, 但是图 7(d)(e)(f)中不管细胞、白细胞还是真菌, 其生成的掩膜边缘都没很好的贴合实际边缘, 相比之下本文方法的图 7(g)(h)(i)掩膜边缘更贴合目标的实际边缘。但是两种方法在细胞重叠处的分割效果都不好。

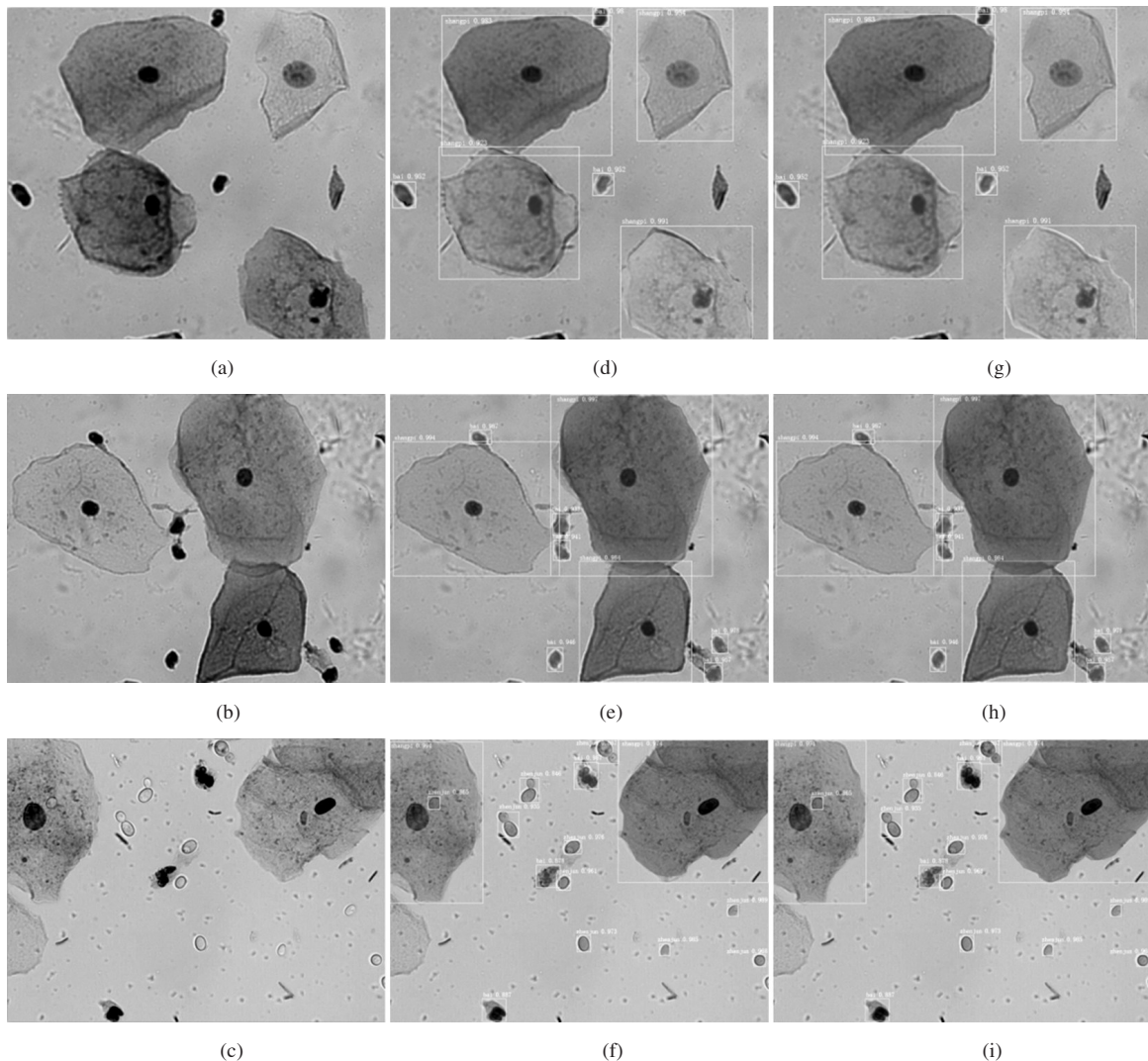


图 7 原图与 Mask R-CNN 效果图

IoU 阈值超过 0.5 时的指标定义为 AP^{50} , 阈值超过 0.75 时为 AP^{75} , 结果见表 2。

表 2 Mask R-CNN 与本文方法测试结果

算法	骨干网络	mIoU	AP^{50}	AP^{75}
Mask R-CNN	ResNet101+FPN	37.3	60.0	39.6
本文方法	ResNet101+DFPN	41.2	63.4	43.1

由表 2 知, 使用 DFPN 代替 Mask R-CNN 中的 FPN 后 mIoU 提高 3.9%。从效果图和数据来看, 本文方法确实对分割效果有所提高。

3 结束语

本文制作了 TCTCOCO 数据集; 实现了 Mask R-CNN 对宫颈细胞图像的实例分割; 提出 DFPN 并且完成了对 Mask R-CNN 的改进, 从实验结果看, 改进后分割精度得到了提高。未来计划对细胞重叠部分进行分割和检测研究, 让细胞重叠部分分割效果更好, 让重叠度较高的框不会被误删。

参考文献(References):

- [1] 万卫兵, 施鹏飞. Snake 活动轮廓在组织培养细胞分割中的

- 应用[J]. 计算机工程与设计, 2006.21:4153-4156
- [2] 胡树煜. 医学图像中粘连细胞分割方法研究[J]. 计算机仿真, 2012.29(2):260-262,27
- [3] 苏士美, 吕雪扬. 骨髓细胞图像的小波变换与K-means聚类分割算法[J]. 郑州大学学报(工学版), 2015.36(4):15-18
- [4] 刻应乾, 曹茂永. 基于Gabor滤波与区域生长的细胞分割[J]. 山东科技大学学报(自然科学版), 2012.31(2):99-103
- [5] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the International Conference on Medical image computing and computer-assisted intervention. Berlin, Germany:Springer, 2015:234-241
- [6] 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. IEEE, 201:779-788
- [7] Liu W, Anguelov D, et al. SSD: Single Shot Multi Box Detector[J]. Computer Vision-ECCV 2016. Springer International Publishing, 2016:21-37
- [8] R. Girshick, Fast R-CNN, in IEEE International Conference on Computer Vision (ICCV), 2015.
- [9] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks[C]. In NIPS, 2015.
- [10] He K, Gkioxari G, Dollar P, et al. Mask r-cnn[C]. International Conference on Computer Vision. New York:IEEE, 2017:2980-2988
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014.39(4):640-651
- [12] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]. Conference on Computer Vision and Pattern Recognition, 2016: 936-944
- [13] Kaiming He, Xiangyu Zhang, et al. Deep Residual Learning for Image Recognition[C]. Conference on Computer Vision and Pattern Recognition, 2015.
- [14] Fisher Yu, Vladlen Koltun. Multi-Scale Context Aggregation by Dilated Convolutions[C]. Conference on Computer Vision and Pattern Recognition, 2016.
- [15] Chen L C, Papandreou G, Kokkinos I, et al. Deep Lab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2016.40(4):834-848
- [16] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context[C]. European Conference on Computer Vision, 2014:740-755



(上接第67页)

- Transformer Networks[J]. 2015. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [7] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2018:7132-7141
- [8] 赵鹏, 刘杨, 刘慧婷等. 基于深度卷积-递归神经网络的手绘草图识别方法[J]. 计算机辅助设计与图形学学报, 2018.2: 217-224
- [9] 刘礼文, 俞强. 循环神经网络(RNN)及应用研究[J]. 科技视界, 2019.32.
- [10] 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
- [11] Liang M, Hu X. [IEEE 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – Boston, MA, USA (2015.6.7–2015.6.12)] 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – Recurrent convolutional neural network for object recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2015:3367-3375
- [12] Lee C Y, Xie S, Gallagher P, et al. Deeply-supervised nets[C]//Artificial intelligence and statistics, 2015:562-570
- [13] Wang F, Jiang M, Qian C, et al. Residual attention network for image classification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:3156-3164
- [14] 袁嘉杰, 张灵, 陈云华. 基于注意力卷积模块的深度神经网络图像识别[J]. 计算机工程与应用, 2019.55(8):9-16
- [15] Srivastava R K, Greff K, Schmidhuber J. Training very deep networks[C]//Advances in neural information processing systems, 2015:2377-2385
- [16] Springenberg J T, Dosovitskiy A, Brox T, et al. Striving for simplicity: The all convolutional net[J]. arXiv preprint arXiv:1412.6806, 2014.

