

一种应用于光遗传激光投影系统的目标检测算法

史再峰^{1*}, 叶鹏¹, 孙诚¹, 罗韬², 王汉杰³, 潘惠卓³

¹天津大学微电子学院, 天津 300072;

²天津大学智能与计算学部, 天津 300072;

³天津大学生命科学学院, 天津 300072

摘要 设计了一种用于无线光传输的激光投影系统并提出了一种基于深度学习的改进型 YOLOv3 (you only look once, v3) 网络用于检测小鼠图像的位置。该网络使用分组卷积对网络参数进行压缩以提高目标检测速度, 使用通道混洗方法以增强网络的信息流通能力。利用交叉熵损失函数中的两个超参数来调整正、负样本的比例以降低易分类样本在损失函数中的权值, 提高了目标检测精度。在 PASCAL VOC2007 和自制小鼠图像数据集上分别进行实验, 结果表明提出的基于改进型 YOLOv3 网络的检测算法检测精度达 90.3%, 检测速度和检测精度都优于传统型网络结构。应用该算法的激光投影系统可以实时检测运动小鼠目标并进行无线光传输等光遗传实验。

关键词 机器视觉; 光遗传; 目标检测; 分组卷积; 通道混洗; 损失函数

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP57.061503

Object Detection Algorithm Applied to Optical Genetic Laser Projection System

Shi Zaifeng^{1*}, Ye Peng¹, Sun Cheng¹, Luo Tao², Wang Hanjie³, Pan Huizhuo³

¹ School of Microelectronics, Tianjin University, Tianjin 300072, China;

² College of Intelligence and Computing, Tianjin University, Tianjin 300072, China;

³ School of Life Sciences, Tianjin University, Tianjin 300072, China

Abstract A laser projection system for wireless light transmission is designed and an modified YOLOv3 (you only look once, v3) network based on deep learning is proposed to detect the location of mouse images. The network first uses packet convolution to compress network parameters to increase target detection speed, and then uses channel shuffle to enhance the network's information flow capabilities. The ratio between the positive sample and the negative sample is adjusted by two hyperparameters on the cross entropy loss function to reduce the weight of the easily classified sample in the loss function, and the detection accuracy is improved. The experimental results on the PASCAL VOC2007 and the self-made mouse image datasets show that the proposed detection algorithm based on the improved YOLOv3 network has a detection accuracy of 90.3%, which is superior to the traditional network structure in terms of detection speed and detection accuracy. The laser projection system using the algorithm can detect moving mouse targets in real time and perform optogenetic experiments such as wireless light transmission.

Key words machine vision; optogenetics; object detection; packet convolution; channel shuffle; loss function

OCIS codes 150.1135; 100.3008; 040.1880

1 引言

目前,光遗传学技术已经成为神经科学研究的重要工具。该技术使用光学信号刺激细胞膜上的光

敏离子通道,从而调控神经细胞活性,对动物活动进行空间和时间上的精确控制^[1-2]。然而,典型的光遗传学实验通常需要在动物(比如小鼠)体内植入光纤进行光传输^[3],在许多行为测试中,这种有线

收稿日期: 2019-07-13; 修回日期: 2019-08-15; 录用日期: 2019-08-28

基金项目: 国家自然科学基金(61674115)

* E-mail: shizaifeng@tju.edu.cn

光源系统阻碍了动物的自由移动,使得动物行为分析复杂化。随着计算机视觉的不断发展,Wang等^[4]提出使用传统的图像检测方法确定动物的位置,配合机械控制系统对其进行无线光传输。传统的图像检测方法受光照等环境因素的影响较大,很容易导致检测失败,而光遗传实验需要对动物进行实时准确的光传输,所以需要一种检测速度极快、检测精度较高、鲁棒性强的目标检测算法用于无线光传输。

近年来,卷积神经网络(CNN)已经成为一种从数据中自动学习特征表示的有效方法^[5],基于CNN的目标检测算法在计算机视觉领域越来越受欢迎^[6]。文献^[7]提出的Faster RCNN(Faster Region Convolutional Neural Networks)采用双步检测策略,将检测过程分为候选区域生成阶段和区域分类阶段,可以大大提高目标的检测精度。但是由于两个阶段需要交互,无法满足实时检测要求。文献^[8]提出了采用单步检测策略的YOLO网络结构,将目标检测问题转化为回归问题,极大地提高了检测速度。文献^[9]提出的SSD(Single Shot multibox Detector)网络结构,结合了YOLO网络的端到端和Faster RCNN的锚点思想来提高物体的检测能力。YOLOv2网络^[10]在YOLO网络的基础上添加了锚点,并使用K-means^[11]算法自动训练边界框以找到更好的目标检测框,从而能够更容易、更准确地检测物体的位置。YOLOv3^[12]网络在Darknet-53的基础上使用了网络适应特性,极大地提高了目标检测精度。

基于CNN的目标检测算法的鲁棒性以及检测精度、速度与传统的目标检测算法相比都有较大提升,而YOLOv3网络是当前目标检测算法满足无线光传输的最优秀算法之一。为了给光遗传实验提供更为稳定的实验环境,本文提出一种基于YOLOv3网络的改进型网络,使用分组卷积的方法减少网络参数,从而减少卷积计算量,加快检测速度。同时,为了提高检测精度,用焦点损失函数替换交叉熵损失函数。为了证明该网络可以应用于光遗传学实验,本文设计了激光投影系统,并采用小鼠图像作为检测对象,制作了小鼠图像数据集。结果表明,改进型网络的检测速度和检测精度都优于YOLOv3网络,完全可以对光遗传学实验中的小鼠图像进行实时目标检测,更好地满足了激光投影系统的需求。

2 基本原理

2.1 激光投影系统结构

为了实现小鼠头部实时、准确的无线光传输,记录小鼠的行走路径、运动体态,以及满足垂直照射小鼠头部的要求,设计了用于光遗传学实验的激光投影系统,如图1所示,其主要由激光投影模块与目标检测模块构成。激光投影模块主要由伺服电机、工字导轨、相机传感器及激光准直器构成。为了使激光投影系统更好地满足无线光传输要求,目标检测模块使用了改进型YOLOv3网络检测小鼠图像的头部位位置并输出位置坐标;激光投影模块接收到位置坐标后,操控伺服电机和激光准直器对小鼠进行无线光传输。

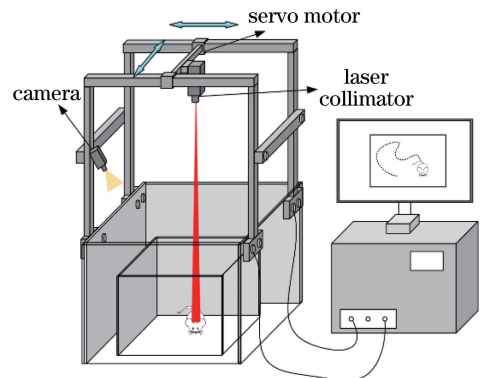


图1 激光投影系统示意图

Fig. 1 Schematic of the laser projection system

激光投影系统使用工字导轨结构,满足了光遗传实验垂直照射小鼠头部的要求;激光准直器和目标检测等模块的配合使用,可以实现对小鼠头部的自动追踪;相机传感器实时记录小鼠的行走路径和运动体态,为光遗传实验的后续分析提供了有效数据。因此,该系统为光遗传实验提供了一种全新的实验环境。

2.2 网络结构

YOLOv3网络由Darknet-53特征提取网络和检测定位网络组成,如图2所示。将图像输入网络后,由Darknet-53对图像进行特征提取,输出三个不同尺度的特征图并进行检测定位。YOLOv3网络采用了残差模块使得网络结构更深,提取特征的能力更强;张量拼接的目的是将Darknet-53中的网络与后面某一特征层的上采样进行拼接,以便进行更好的特征融合,从而提高检测精度。本文主要对YOLOv3网络的Darknet-53特征提取网络和部分

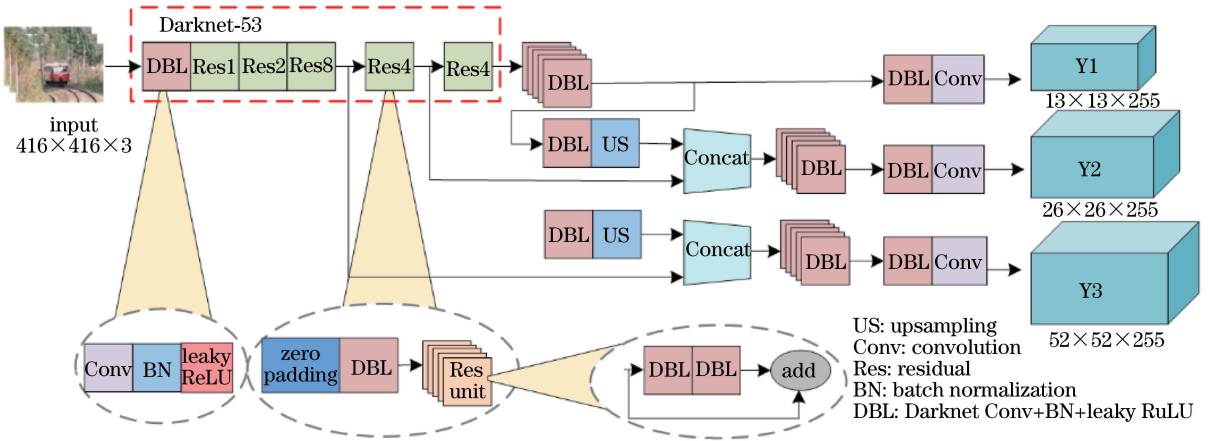


图 2 网络结构图

Fig. 2 Network structure diagram

损失函数进行了改进。

为了压缩网络参数,改进型网络使用分组卷积的方法替换 Darknet-53 特征提取网络中的标准卷积。使用分组卷积^[13]的方法(分为 4 组)进行 3×3 卷积以获得特征图;采用通道混洗^[14]增强网络间的信息流通能力。因为在第一层卷积必须从输入图像中获得尽可能多的特征,所以该卷积层使用标准卷积。

2.3 分组卷积

分组卷积是一种特殊的稀疏连接卷积。分组卷积操作流程如图 3(a)所示,先将输入的特征图在通道维度上进行分组,再将相应的卷积核分成相同的组,最后将各组特征图和卷积核分别进行卷积运算得到输出特征图。假设一个输入特征图的通道深度为 a ,输出特征图的通道深度为 b , k 和 l 为特征图 F 的高和宽(假定两者相同)。用 $i \times i$ 大小的标准卷积核 K 对特征图进行卷积。设步长和填充都为 1,标准卷积的计算量 M 可以表示为

$$M_{k,l,b} = \sum_{i,a} K_{i,i,a,b} \times F_{k+i-1,l+i-1,a} \quad (1)$$

假设分组数为 G ,则分组卷积的计算量 N 可以

表示为

$$N_{k,l,b} = \sum_{i,a} K_{i,i,a/G,b} \times F_{k+i-1,l+i-1,a/G} \times G \quad (2)$$

分解卷积与标准卷积的计算量比例关系式可以表示为

$$\frac{\sum_{i,a} K_{i,i,a/G,b} \times F_{k+i-1,l+i-1,a/G} \times G}{\sum_{i,a} K_{i,i,a,b} \times F_{k+i-1,l+i-1,a}} = 1/G \quad (3)$$

从(3)式可以得到,分组卷积的卷积计算量是标准卷积的 $1/G$,减少的计算量与分组数成正比。

在分组卷积降低网络卷积的计算量的同时,由于将特征图进行了分组,这会导致各组之间的特征图存在信息不畅通的问题,这无疑会影响网络的检测精度。为了解决这一问题,采用通道混洗的方法。将各组特征图的通道进行切割重组,可以增强网络之间的信息流通,提高信息表达能力,通道混洗原理如图 3(b)所示。先将分为 G 组的特征图进行变形,再将变形后的参数矩阵 G_1 等进行转置操作,最后将各组参数矩阵作平坦化操作,实现各组特征图在通道维度上的混合。

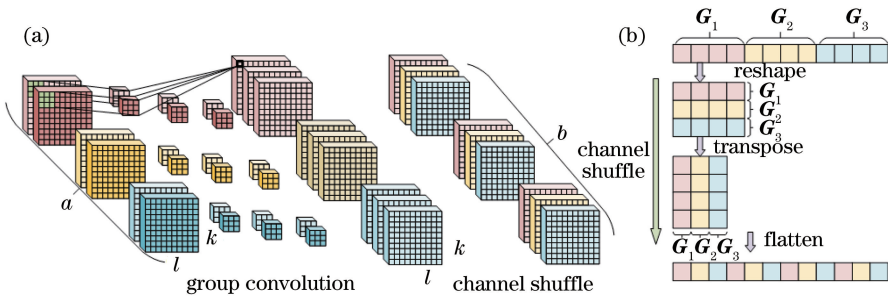


图 3 分组卷积。(a)操作流程;(b)通道混洗

Fig. 3 Packet convolution. (a) Operation flow; (b) channel shuffle

2.4 损失函数

评价一种目标检测算法的优劣,除了检测速度要求外,还有一个重要指标,即检测精度。本文网络是基于 YOLOv3 网络改进的目标检测方法,属于单步检测方法。文献[15]指出单步检测方法的检测精度不如双步检测方法,原因在于没有很好地解决正负样本分布不均衡的问题。本文网络中采用了焦点损失函数,在交叉熵损失函数的基础上增加两个超参数,可解决分布不均衡的问题。对于二分类,标准的交叉熵损失函数(C)为

$$C_{(p,y)} = \begin{cases} -\log p & y=1 \\ -\log(1-p) & \text{otherwise} \end{cases}, \quad (4)$$

式中: $y \in \{-1, 1\}$ 为真实样本的标签; $p \in [0, 1]$ 为模型对于标签 $y=1$ 的估计概率。文献[15]指出,使用交叉熵损失函数时,由于负样本的数量远大于正样本,而且都容易分类,所以在代价函数中占有绝对的比重。如果不加以控制,很容易导致网络在训练过程中出现恶化。焦点损失函数可以有效解决这一问题,使得单步检测方法的检测精度得以改善。焦点损失函数(F)定义为

$$F_{(p_i)} = -\alpha_i (1-p_i)^\gamma \log(p_i), \quad (5)$$

式中: p_i 为不同类别的分类概率; γ 是大于 0 的值; α_i 是属于 $[0, 1]$ 区间的小数。 γ 和 α_i 都是超参数,不参与训练。由(5)式可以看到无论是正样本还是负样本, p_i 值越大,权重 $(1-p_i)^\gamma$ 值就越小,说明易分类样本的比重可以通过 γ 的取值进行抑制,这就增加了难分类样本的重要性。 α_i 值用于调节正样本和负样本的比例,正样本使用 α_i 时,对应的负样本使用 $1-\alpha_i$ 。 γ 和 α_i 的最优值是相互影响的,所以在评估检测精度时需要把两者组合起来进行调节。

3 分析与讨论

3.1 实验准备

实验需要的小鼠图像数据集委托天津大学生命科学学院进行采集。考虑到实验环境影响检测性能,在图像采集的过程中使用了多角度拍摄以及不同的照明条件。实验一共采集了 1500 张形态各异、不同颜色的小鼠图像。为了增强实验数据集的丰富性,将收集的图像在颜色、亮度、旋转等方面进行了图像预处理,增强后的图片数量达到 15000 张,并按照 8:2 的比例分为训练集和测试集,部分图像数据集图片如图 4 所示。为了配合改进型 YOLO v3 网络,将输入图像的分辨率调整为 $416 \text{ pixel} \times$

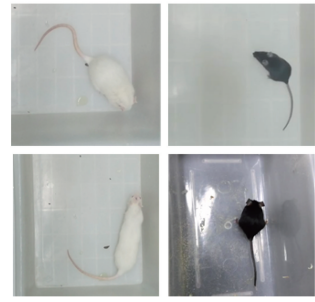


图 4 部分小鼠图像数据集图片

Fig. 4 Part of the mouse image dataset

416 pixel。光遗传实验通常在小鼠头部接入光纤进行光传输,所以对图像中的小鼠头部进行手动标注并进行分类。

实验中网络搭建和训练环境如下:GPU 型号为 NVIDIA RTX2080,64 位 Windows10 操作系统以及 TensorFlow 1.8.0 版本。训练迭代次数调整为 70000 次,初始学习率、权重衰减正则化参数等与 YOLOv3 网络设置一致,学习率在 40000 步后降至 0.0001,在 50000 步后降至 0.00001。

3.2 检测实验

焦点损失函数中有两个超参数 γ 和 α_i ,其值直接影响网络的性能。因此,实验首先需要确定 γ 和 α_i 的值。超参数使用不同值时,改进型网络在测试集中取得的均值平均精度(mAP)如图 5 所示。由图 5 可以看到,当 $\gamma=2$ 且 $\alpha_i=0.25$ 时,网络的检测精度最高。

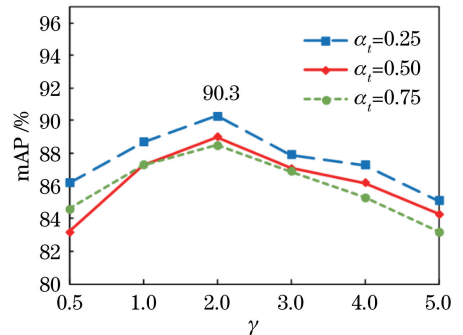


图 5 不同 γ 和 α_i 值下的 mAP

Fig. 5 mAP at different γ and α_i values

改进型 YOLOv3 网络、SSD 网络及 YOLOv3 网络在小鼠图像数据集的准确率(Precision)与召回率(Recall)结果曲线,如图 6 所示。由图 6 可以看到,改进型 YOLOv3 网络在自制小鼠图像数据集上的准确率与召回率都高于 YOLOv3 和 SSD 目标检测网络。

不同主流检测网络及改进型 YOLOv3 网络在

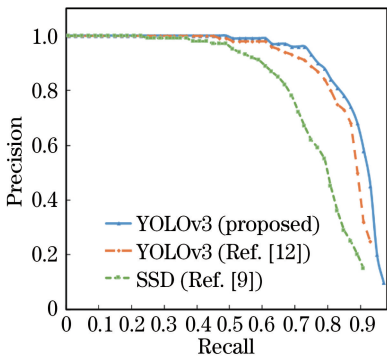


图 6 不同算法在老鼠图像数据集上的召回率与准确率结果曲线

Fig. 6 Recall and precision results of different algorithms on mouse image dataset

老鼠图像数据集及 PASCAL VOC2007^[16]数据集上的 mAP 值和检测时间如表 1 所示。Faster RCNN 在老鼠图像数据集的 mAP 值达到 91.4%，但是检

表 1 不同算法的精度与速度结果

Table 1 mAP and speed results of different algorithms

Method	PASCAL VOC2007		Mouse image dataset	
	mAP / %	Speed / ms	mAP / %	Speed / ms
Faster RCNN ^[7]	73.2	200	91.4	208
SSD ^[9]	74.3	61	79.3	64
YOLOv3 ^[12]	79.2	29	88.4	32
YOLOv3 (proposed)	79.8	25	90.3	26

不同算法的老鼠图像检测对比如图 7 所示。由图 7 可以看到,改进后的网络对黑色小鼠和白色小鼠的检测框的置信度低于 Faster RCNN,但是高于 SSD 及 YOLOv3 网络,说明该网络对老鼠头部图像的检测精度较高。同时,改进型 YOLOv3 网络检测老鼠图像时,有极少检测失败的情况出现,这是网络应进一步优化的方向。

测速度无法满足实时检测的要求;SSD 网络的检测速度显著提升,但与其他网络算法相比 mAP 值较低;改进型 YOLOv3 检测网络在老鼠图像数据集的 mAP 值高达 90.3%,虽然低于 Faster RCNN,但是检测速度约为 Faster RCNN 的 8 倍。

由图 6 和表 1 可以看到,将分组卷积和焦点损失函数应用于 YOLOv3 网络,可以很好地提高网络的检测速度和检测精度。改进型 YOLOv3 网络在 PASCAL VOC2007 和老鼠图像数据集上的 mAP 值分别为 79.8%和 90.3%,每个数据集图像的平均处理时间仅为 25 ms 和 26 ms。根据目前的实验器材,激光投影系统对目标检测算法 mAP 值的要求是 85% 以上,检测时间是 60 ms 以内。Faster RCNN 和 SSD 网络无法满足该要求,而改进型 YOLOv3 网络比 YOLOv3 网络的 mAP 值更高,检测速度更快,更适合应用于激光投影系统。

4 结 论

结合 YOLOv3 网络出色的目标检测实时性与准确性,提出一种改进型 YOLOv3 目标检测网络用于辅助光遗传实验。为了提高检测速度与检测精度,使用了分组卷积和改进交叉熵损失函数的方法。实验结果表明,改进后的网络对老鼠头部图像的检

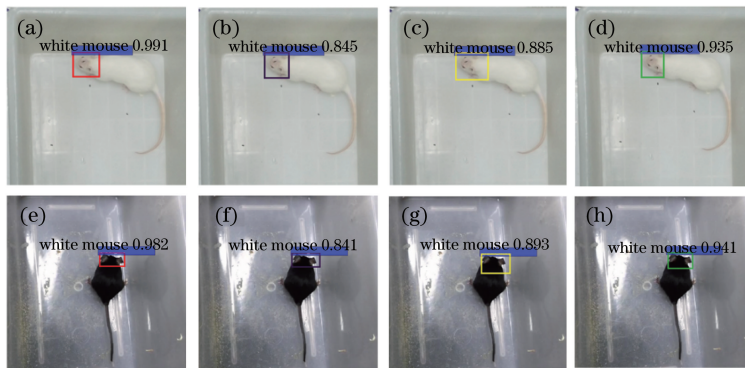


图 7 不同算法的老鼠图像检测对比。(a)(e) Faster RCNN; (b)(f) SSD; (c)(g) YOLOv3; (d)(h)改进型 YOLOv3

Fig. 7 Comparison of mouse image detection with different algorithms. (a)(e) Faster RCNN; (b)(f) SSD; (c)(g) YOLOv3; (d)(h) proposed YOLOv3

测精度达到 90.3%，检测速度达到 26 ms，优于传统的 YOLOv3 网络，完全可以满足激光投影系统对小鼠进行实时检测的要求。提出的激光投影系统可以为光遗传实验提供稳定可靠的无线光传输环境，使小鼠可以自由移动，并且进一步改善脑部神经刺激范围，为远程控制小鼠脑部的神经活动开辟了新的可能性。

参 考 文 献

- [1] Yizhar O, Fenno L E, Davidson T J, et al. Optogenetics in neural systems[J]. *Neuron*, 2011, 71(1): 9-34.
- [2] Boyden E S, Zhang F, Bamberg E, et al. Millisecond-timescale, genetically targeted optical control of neural activity[J]. *Nature Neuroscience*, 2005, 8(9): 1263-1268.
- [3] Zhang F, Gradinaru V, Adamantidis A R, et al. Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures[J]. *Nature Protocols*, 2010, 5(3): 439-456.
- [4] Wang Y, Lin X D, Chen X, et al. Tetherless near-infrared control of brain activity in behaving animals using fully implantable upconversion microdevices[J]. *Biomaterials*, 2017, 142: 136-148.
- [5] Long X, Su H S, Liu G H, et al. A face recognition algorithm based on angular distance loss function and convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2018, 55(12): 121505.
龙鑫, 苏寒松, 刘高华, 等. 一种基于角度距离损失函数和卷积神经网络的人脸识别算法[J]. *激光与光电子学进展*, 2018, 55(12): 121505.
- [6] Ou P, Zhang Z, Lu K, et al. Object detection in of remote sensing images based on convolutional neural networks [J]. *Laser & Optoelectronics Progress*, 2019, 56(5): 051002.
欧攀, 张正, 路奎, 等. 基于卷积神经网络的遥感图像目标检测[J]. *激光与光电子学进展*, 2019, 56(5): 051002.
- [7] 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.
- [8] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6517-6525.
- [11] Tang X L, Yang W, Hu X S, et al. A novel simplified model for torsional vibration analysis of a series-parallel hybrid electric vehicle[J]. *Mechanical Systems and Signal Processing*, 2017, 85: 329-338.
- [12] Redmon J, Farhadi A. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2018-09-07]. <https://arxiv.org/abs/1804.02767>.
- [13] Huang G, Liu S C, van der Maaten L, et al. CondenseNet: an efficient DenseNet using learned group convolutions[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2752-2761.
- [14] Zhang X Y, Zhou X Y, Lin M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 6848-6856.
- [15] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018: 1.
- [16] Everingham M, van 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.