

基于GAM-YOLOv8算法的生活垃圾检测

陈思羽¹, 舒航², 陈宇阳¹, 王晓峰¹, 李海霞¹, 孙贾梦²

¹重庆科技大学数理与大数据学院, 重庆

²重庆科技大学电气工程学院, 重庆

收稿日期: 2024年3月19日; 录用日期: 2024年4月30日; 发布日期: 2024年5月9日

摘要

日常生活垃圾分拣是困扰人们的一个难题, 生活垃圾的种类繁多, 所处环境复杂, 常用的目标检测算法无法适应各种复杂的环境, 导致精度较低。为了准确的分拣生活垃圾, 提出了一种基于GAM注意力机制的YOLOv8生活垃圾检测算法。该算法在YOLOv8优秀的目标检测基础上, 加入GAM注意力机制, 增强网络对重要通道特征信息的关注能力, 提升高层网络中图像特征语义信息的提取能力, 提高复杂环境垃圾分类检测精度的效果。实验表明, 在40多种生活垃圾类别检测测试中, 改进的YOLOv8算法mAP平均精度84.5%, 较原始算法YOLOv8提升了0.7%。因此改进的YOLOv8算法可以通过对垃圾图像的分析 and 识别, 帮助人们准确地进行垃圾分类。较好的满足了生活垃圾检测精度的要求。

关键词

YOLOv8, 生活垃圾检测, GAM注意力机制, 增强网络

Domestic Waste Detection Based on GAM-YOLOv8 Algorithm

Siyu Chen¹, Hang Shu², Yuyang Chen¹, Xiaofeng Wang¹, Haixia Li¹, Jiameng Sun²

¹College of Mathematics and Big Data, Chongqing University of Science and Technology, Chongqing

²School of Electrical Engineering, Chongqing University of Science and Technology, Chongqing

Received: Mar. 19th, 2024; accepted: Apr. 30th, 2024; published: May 9th, 2024

Abstract

The sorting of daily household garbage is a difficult problem that troubles people. There are many kinds of household garbage and the environment is complicated. In order to sort domestic waste

文章引用: 陈思羽, 舒航, 陈宇阳, 王晓峰, 李海霞, 孙贾梦. 基于 GAM-YOLOv8 算法的生活垃圾检测[J]. 人工智能与机器人研究, 2024, 13(2): 194-202. DOI: 10.12677/airr.2024.132021

accurately, a YOLOv8 domestic waste detection algorithm based on GAM attention mechanism was proposed. Based on the excellent target detection of YOLOv8, this algorithm adds GAM attention mechanism to enhance the network's ability to pay attention to important channel feature information and improve the extraction ability of image feature semantic information in high-level networks. Experiments show that in more than 40 types of domestic waste detection tests, the average accuracy of the improved YOLOv8 algorithm mAP is 84.5%, which is 0.7% higher than that of the original YOLOv8 algorithm. Therefore, the improved YOLOv8 algorithm can help people to classify garbage quickly and accurately through the analysis and recognition of garbage images. It can meet the requirement of detecting precision of domestic waste.

Keywords

YOLOv8, Domestic Waste Detection, GAM Attention Mechanism, Enhanced Network

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

生活垃圾分类作为一项重要的环保举措，旨在降低垃圾对环境造成的危害，促进资源的循环利用。随着社会的发展和城市化进程的加快，生活垃圾的数量不断增加，垃圾处理难题逐渐凸显。正确的生活垃圾分类可以有效地减少垃圾填埋量和焚烧产生的二氧化碳排放，降低对环境的负面影响。因此，各国政府和社会各界都强调生活垃圾分类的重要性。

目标检测算法作为计算机视觉领域的一个重要分支在物体检测、分类上具有优异表现。同时，生活垃圾检测任务需要对垃圾的位置和类别进行精准计算与辨别区分。目标检测算法是计算机视觉领域的一项重要技术，旨在识别和定位图像或视频中的不同目标物体。随着人工智能和深度学习的快速发展，目标检测算法的准确度和效率不断提升，为各种应用场景提供了广阔的可能性。诸如基于区域的方法(如 R-CNN [1]、Fast R-CNN [2])、基于单阶段的方法(如 YOLO [3]、SSD [4]、EfficientDet [5])等多种目标检测算法得到了广泛应用。这些算法在准确度、速度、鲁棒性和适应性等方面都有着一定的优势和特点。当前，以 SSD、YOLO 系列为代表的一阶段目标检测算法，能够较好的避免二阶段目标检测算法存在的问题。通过端到端的方式直接利用卷积神经网络实现回归预测和分类，从而获得物体的位置和类别信息。目前，大多数一阶段目标检测算法也都存在着对特征通道重要信息的关注力度较弱的问题。

针对上述目标检测算法在生活垃圾检测上存在的问题，本文在 YOLOv8 [6]目标检测算法的基础上提出了一种基于 GAM 注意力机制[7]的改进 YOLOv8 生活垃圾检测算法。在原始 YOLOv8 模型的主干特征提取网络中设计了 GAM 通道注意力机制，自适应地调整各特征通道响应值，强化对生活垃圾图像重要特征通道信息的提取能力。

2. 改进 YOLOv8 算法介绍

2.1. YOLOv8 算法介绍

YOLO 是一种基于图像全局信息进行预测的目标检测系统。自 2015 年 Joseph Redmon、Ali Farhadi 等人提出初代模型以来，领域内的研究者们已经对 YOLO 进行了多次更新迭代，模型性能越来越强大。

现在, YOLOv8 已正式发布。YOLOv8 算法是一种快速的单阶段目标检测方法, 包括输入段、主干、颈部和输出段。输入段对输入图像执行镶嵌数据增强、自适应锚计算和自适应灰度填充。主干网络和颈部模块构成 YOLOv8 网络的中心结构。输入图像由多个 Conv 和 C2f 模块处理, 以提取不同尺度的特征图。C2f 模块是原始 C3 模块的改进版本, 用作主要的残差学习模块。它结合了 YOLOv7 [8] 中 ELAN 结构的优点, 减少了一个标准卷积层, 并充分利用瓶颈模块来增强梯度分支。该方法不仅保留了梯度流的轻量化特征, 而且能够捕获更丰富的梯度流信息。图 1 描述了算法的基本结构。输出的特征图由 SPPF 模块处理, 该模块采用具有不同内核大小的池化来联合收割机特征图, 然后将结果传递到颈部层。YOLOv8 的颈部层结合了 FPN [9] + PAN [10] 结构, 以增强模型的特征融合能力。该结构使用上采样和下采样技术将高级和低级特征映射相结合, 促进了语义和本地化特征的传递。通过这种方法, 网络变得更好地融合来自对象的特征。YOLOv8 的网络结构如图 1 所示。

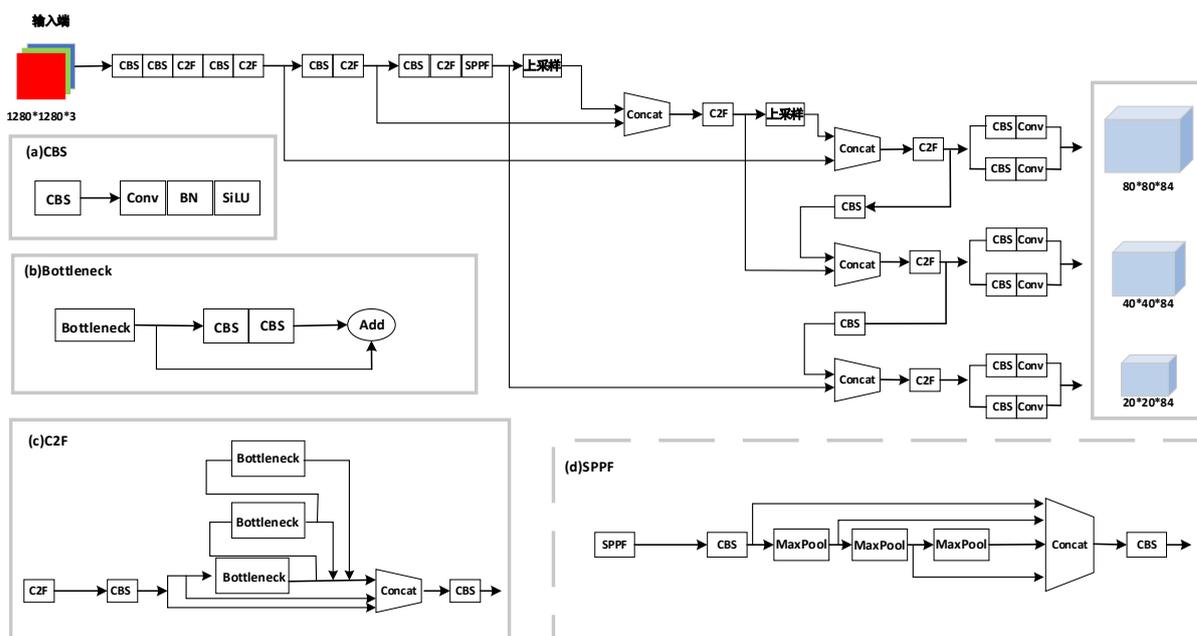


Figure 1. YOLOv8 Network structure
图 1. YOLOv8 网络结构

2.2. GAM 注意力机制介绍

GAM (Global Attention Mechanism) 是一种注意力机制, 用于加强神经网络对输入数据的关注程度。它通过将全局的上下文信息与局部特征进行交互, 提高模型的表达能力和泛化能力。

GAM 的基本思想是在神经网络的不同层之间引入注意力模块, 使网络能够自适应地选择输入数据中重要的特征。具体而言, GAM 综合了全局平均池化和全局最大池化的结果, 生成一个全局向量, 用于表示输入特征的全局信息。然后, 它将全局向量与局部特征进行加权相加或拼接, 以产生更丰富的特征表示。

GAM 主要有两个关键组成部分: 自适应的全局池化和特征融合。自适应的全局池化通过学习全局权重的方式, 将输入特征进行加权平均, 得到全局向量。特征融合阶段, 将全局向量与局部特征进行融合, 通常有加权相加或拼接两种方式。这样, GAM 能够自动学习到不同尺度的特征之间的关系, 增强网络对全局信息的感知和关注。

GAM 的引入可以提升模型的性能，并且具有很好的可解释性。它能够帮助网络更好地理解输入数据的上下文信息，对重要特征的关注更加集中，从而改善模型的分类、回归或其他任务的性能。GAM 在图像识别、目标检测和自然语言处理等任务中都得到了广泛应用，并在提升模型性能方面取得了显著效果。其结构图如图 2 所示。

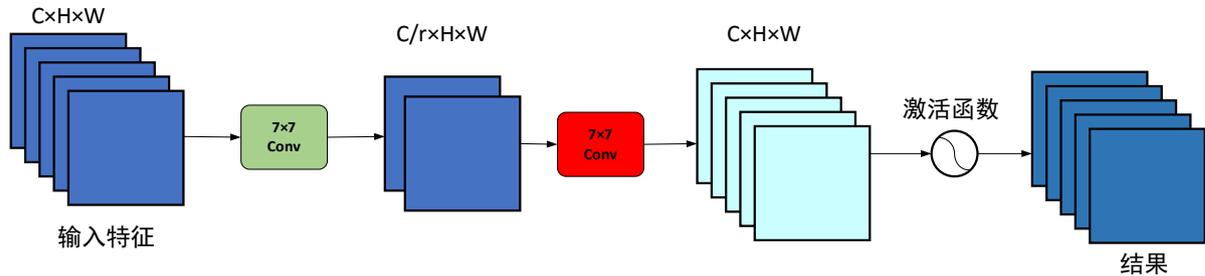


Figure 2. GAM attention mechanism
图 2. GAM 注意力机制

2.3. 改进的 YOLOv8 模型设计

在原始 YOLOv8 算法主干特征提取网络中有 6 个 C2f 结构，通过相互堆叠的方式实现对图像特征的提取，主干特征提取网络见图 3。由于生活垃圾图像特征的多样性，原始 YOLOv8 算法的主干特征提取网络采用均衡特征通道信息的方式，但这样的方式将导致主要与次要通道特征信息区分能力的缺失，导致在针对生活垃圾图像多样性的特征信息提取能力上表现较弱。而 GAM 通道注意力机制充分利用不同特征通道信息重要程度的权重，给输入的特征信息施加权重，增强重要特征通道信息获取能力，抑制不重要特征通道信息。本文通过在主干特征提取网络中引入注意力机制的方式来加强其对多样性的生活垃圾图像特征信息的提取性能。改进的网络结构如图 3 所示。

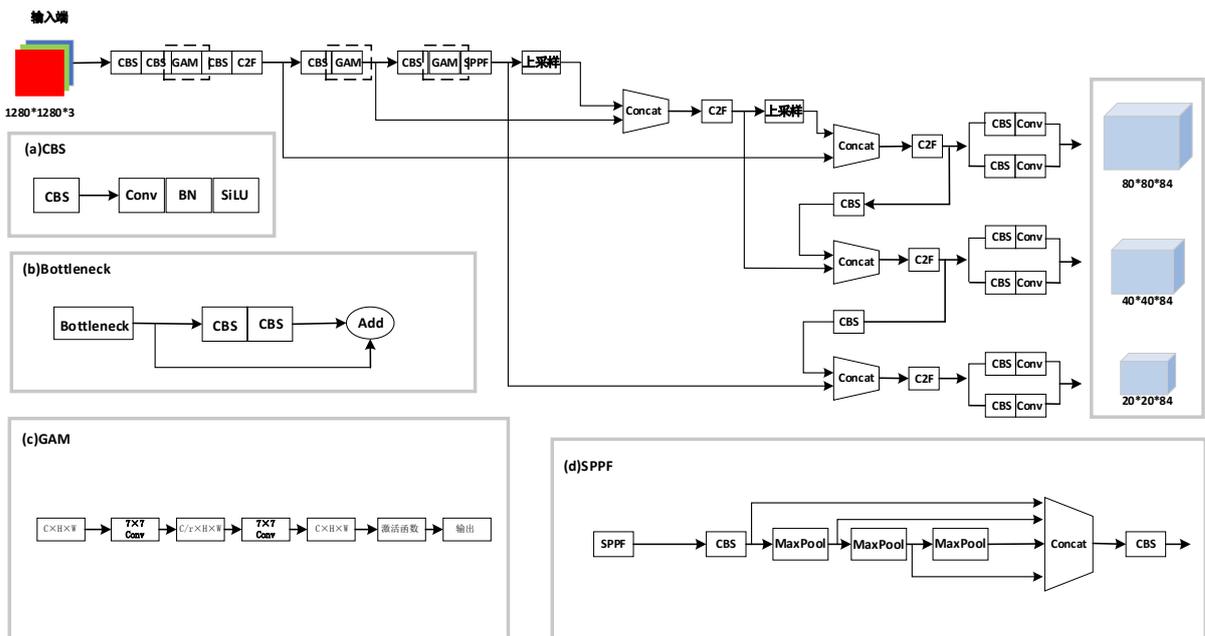


Figure 3. GAM-YOLOv8 Network structure
图 3. GAM-YOLOv8 网络结构

3. 相关工作

3.1. 实验环境

本文提出的生活垃圾图像检测算法是基于 Pytorch1.7 框架及 Python3.10 版本下实现的,在 Windows11 系统下使用 Intel Core i7-11700F 2.50 GHZ 处理器、内存 32 GB, 显卡为 NVIDIA GeForce RTX4060, 显卡内存 12 GB。

3.2. 数据集

本文选用了“华为云杯”2020 深圳开放数据应用创新大赛·生活垃圾图片分类初赛图像数据作为研究对象。在对数据进行充分预处理后,我们从数据集中剔除了不具代表性的生活垃圾图像,精选出了包括果皮果肉、易拉罐、食用油桶等在内的 40 多种生活垃圾图像,共计 10,000 张,作为本文研究的生活垃圾检测数据集(见图 4)。为了确保研究模型的训练和测试有效性,我们将数据按照 9:1 的比例进行划分,其中 9000 张用于训练,1000 张用于测试。同时,我们对图像进行了统一的裁剪以适应分辨率。此外,我们还采用了随机旋转、平移、缩放等数据增强处理方法,并结合 Mosaic 数据增强技术,以提升数据的复杂性,从而提高模型的泛化性能。这些步骤和方法的综合应用,为我们的研究提供了更为充分和准确的数据基础,为模型训练和性能优化提供了有力支持。数据集网址: <https://www.huaweicloud.com/>。

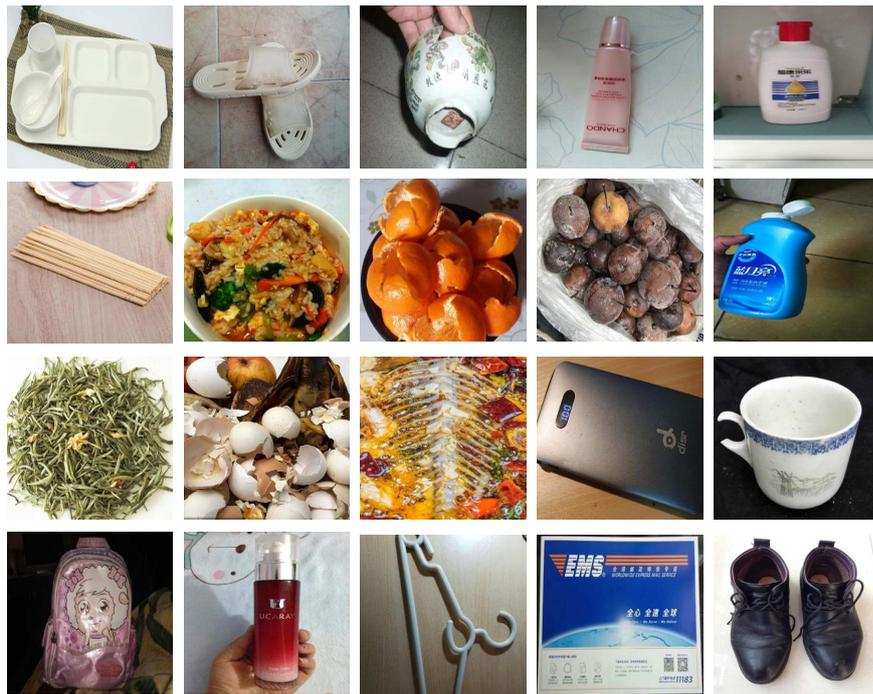


Figure 4. Domestic waste data set
图 4. 生活垃圾数据集

4. 改进 YOLOv8 算法介绍

4.1. 评价指标

根据目标检测领域常用的评估指标,生活垃圾分类检测效果可以通过精确率(precision)、召回率(recall)和平均精度均值(mAP)这三个指标来评估其好坏。

式(1)是精确率和召回率的计算公式:

$$\begin{cases} \text{precision} = \frac{t_p}{t_p + f_p} \\ \text{recall} = \frac{t_p}{t_p + f_n} \end{cases} \quad (1)$$

其中 t_p 表示实际为正样本, 且划分正确的数量; f_n 表示实际为负样本, 且划分正确的数量; f_p 表示实际为正样本, 但被分为负样本的数量。

mAP 的计算公式如式(2)所示。

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AveP}(i) \quad (2)$$

其中 N 为类别数; $\text{AveP}(i)$ 为第 i 类的平均精确率。

4.2. 性能对比

为了量化本文改进的 YOLOv8 算法的性能表现, 将 YOLOv8 检测算法在日常生活垃圾目标检测任务上的精确率(precision)、召回率(recall)和平均精度均值(mAP)作对比。它们是用来衡量模型性能的重要指标。精确率指标表示了模型在预测为正例的样本中, 实际为正例的比例, 而召回率则表示了模型能够正确预测为正例的样本在所有正例样本中的比例。平均精度均值是目标检测任务中用来综合考虑精确率和召回率的指标, 对于多类别目标检测任务来说尤为重要, 它代表了所有类别的平均精度值的均值。在这些指标中, 通常情况下, 高精确率和高召回率是期望的, 而对于 mAP 来说, 高值表示模型在多类别目标检测任务中的整体性能较好。对比结果如表 1 所示。

Table 1. Performance comparison of YOLOv8 model before and after improvement

表 1. YOLOv8 模型改进前后性能对比

主干特征提取网络	精度/%	召回率/%	平均精度均值/%
YOLOv8 [6]	86.49	79.54	83.83
GAM-YOLOv8	89.36	83.46	84.53





(a) YOLOv8

(b) YOLOv8 改进算法

Figure 5. Comparison of simple scene results
图 5. 简单场景结果对比



Figure 6. Comparison of complex scene results
图 6. 复杂场景结果对比

4.3. 结果对比

本文选择 YOLO 系列最新版本的 YOLOv8 算法, 对日常生活垃圾进行分类识别, 在 YOLOv8 的基础上对其进行了改进。处理结果得到了一定的提升, 我们选择了其中几组对比展示。如图 5(a), 图 5(b) 所示。

由两组对比图可以看出, 在简单的检测场景中结合 GAM 注意力机制的 YOLOv8 网络较原来算法的检测精度有一定提升。复杂场景的结果对比如图 6(a), 图 6(b) 所示。

由两组对比图可以看出, 在复杂的检测场景中结合 GAM 注意力机制的 YOLOv8 网络较原来算法的检测精度也有一定提升。综上所述, 本文改进算法对垃圾检测的精度有一定的提升。

5. 结论

本文针对生活垃圾检测领域面临的多样性、复杂性和难度大的问题进行了深入研究。我们提出了一种改进的 YOLOv8 生活垃圾目标检测算法, 以有效克服主干特征提取网络在获取图像通道特征信息方面的限制, 并加强通道注意力特征提取能力, 从而解决了正负样本以及难易样本分布不均衡的问题。

实验结果明确显示, 我们的算法在 40 多个生活垃圾类别的分类识别上相较于原始 YOLOv8 模型取得了显著提升, 平均精度提高了 0.70%。尤其值得注意的是, 在保持检测速度几乎不变的情况下, 我们的算法能够更好地适应生活场景中垃圾检测对于高精度要求的需求。这不仅为生活垃圾管理提供了更可靠、高效的解决方案, 同时也为目标检测领域的进一步研究提供了有价值的经验和方法。

基金项目

国家大学生科技创新训练项目(202311551016)。

参考文献

- [1] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [2] 马雯, 于炯, 王潇, 陈嘉颖. 基于改进 Faster R-CNN 的垃圾检测与分类方法[J]. 计算机工程, 2021, 47(8): 294-300.
- [3] 邵延华, 张铎, 楚红雨, 张晓强, 饶云波. 基于深度学习的 YOLO 目标检测综述[J]. 电子与信息学报, 2022, 35(16): 156-162.
- [4] 赵珊, 刘子路, 郑爱玲, 高雨. 基于 MobileNetV2 和 IFPN 改进的 SSD 垃圾实时分类检测方法[J]. 计算机应用, 2022, 42(S1): 106-111.
- [5] Tan, M.X., Pang, R.M. and Le, Q.V. (2020) Efficient Det: Scalable and Efficient Object Detection. *CVPR*, **26**, 125-163.
- [6] Talaat, F.M. and Eldin, H.Z. (2023) An Improved Fire Detection Approach Based on YOLO-v8 for Smart Cities. *Neural Computing and Applications*, **35**, 20939-20954. <https://doi.org/10.1007/s00521-023-08809-1>
- [7] 袁磊, 唐海, 陈彦蓉, 等. 改进 YOLOv5 的复杂环境道路目标检测方法[J]. 计算机工程与应用, 2023, 59(16): 268-293.
- [8] Wang, C.-Y., Bochkovskiy, A. and Liao, H.-Y.M. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, BC, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/CVPR52729.2023.00721>
- [9] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 21-26 July 2017, 2117-2125. <https://doi.org/10.1109/CVPR.2017.106>
- [10] Li, H., Xiong, P., An, J. and Wang, L. (2018) Pyramid Attention Network for Semantic Segmentation. <http://arxiv.org/abs/1805.10180>