

基于 YOLOv8 的蛇类识别系统

崔祥辉 张锦雄* 李陶深

广西大学计算机与电子信息学院, 南宁 530004

摘要 准确快速识别蛇类及其毒性对于蛇伤患者的抢救至关重要, 蛇类数量及种群分布的了解是蛇类有效保护和蛇类医药保健价值充分利用的前提。本文提出了一种基于深度学习模型 YOLOv8 的蛇类识别方法, 利用 Pyside6 构建设计实现了一个蛇类识别系统。该系统集成了数据采集、预处理、建模以及识别结果展示等多种技术, 实现对蛇类图像的数据清洗、数据标注和数据增强预处理; 运用 PythonGUI 图形界面 Pyside6 开发了系统展示界面, 实现前后端稳定适配的连接。系统性能测试结果表明, 该系统提供了图片检测、摄像头检测、视频检测等功能, 能够为用户提供准确快速、实时和高效的蛇类识别服务。

关键字 图像识别, 蛇类识别系统, 卷积神经网络, YOLOv8

YOLOv8-based Snake Recognition System

Cui XiangHui Zhang JinXiong Li TaoShen

School of Computer of Electronics and Information of GuangXi University
Nanning 530004, China

2791839363@qq.com zhangjx@gxu.edu.cn tshli@gxu.edu.cn

Abstract—Accurately and quickly identifying snakes and their toxicity is crucial for the rescue of snake injured patients. Understanding the number and population distribution of snakes is a prerequisite for effective protection of snakes and the full utilization of their medical and health value. This paper proposes a snake recognition method based on the deep learning model YOLOv8, and designs and implements a snake recognition system using pyside6. The system integrates various technologies such as data acquisition, preprocessing, modeling, and recognition result display to achieve data cleaning, data annotation, and data enhancement preprocessing of snake images. The system display interface is developed using the Python GUI graphical interface (Pyside6) to perform stable and adaptable connections between the front-end and back-end. The system performance test results show that the system provides functions such as image detection, camera detection, and video detection, which can provide users with accurate, fast, real-time, and efficient snake recognition services.

Keywords—Image recognition, Snake recognition system, Convolutional neural networks, YOLOv8

1 引言

蛇作为一种爬行动物, 是大自然中不可或缺的一部分。我国蛇类资源十分丰富, 拥有 200 多种蛇类, 其中约 88% 蛇类分布在我国南方地区^[1]。蛇类可分为无毒蛇和有毒蛇两种。毒蛇伤人事件时有发生, 患者被毒蛇咬伤后需要根据其种类选择对应的抗蛇毒血清进行治疗, 否则会有生命之虞。另外, 某些种类的蛇在医药、保健等领域有一定的经济价值。因此, 蛇类的识别具有重要的现实意义。此外, 蛇类的识别可以让人类更好的了解蛇的分布、数量和环境需求, 从而能够制定有效的蛇类保护措施, 以保持地球生物多样性^[2]。

世界上现有二千七百多种蛇, 分属十科, 有毒蛇约有四百二十种, 约占总数的百分之十五^[1]。由于蛇

种类繁多, 准确识别蛇类, 尤其准确识别毒蛇的种类对非专业人士形成了严峻的挑战。

2 相关工作

近年来, 深度学习技术发展飞快, 卷积神经网络等深度学习技术在计算机视觉和自然语言处理等领域都得到了应用。

2.1 基于深度学习的蛇类识别

同样, 基于深度学习的图像识别技术在蛇类识别问题上也得到了应用^[3-6]。张^[3]基于迁移学习方法重新训练 Inception-v3 模型对 5 种具代表性蛇类进行识别, 并在 Web 上完成应用部署。付^[4]首先建立 10 种中国常见蛇类共计 10336 幅图像的 CHINESESNAKES 蛇类图像数据集, 然后设计 BRC 卷积神经网络以识别蛇类, 最后设计实现一个基于安卓的蛇类图像识别

*通讯作者: 张锦雄, 邮箱: zhangjx@gxu.edu.cn

系统。鲁和罗^[5]运用轻量级网络结构 Squeezenet 模型对 10 种蛇类进行图像识别,在保证识别精度的同时很好地满足蛇类识别应用的可靠性和实时性。周等人^[6]以利用 ImageNet 数据集训练结果作为基础进行模型初始化,采用迁移学习方式,运用 Adam 算法优化训练,有效提升蛇类识别任务的准确度。

2.2 YOLO 算法系列

YOLO^[7]算法是一种基于深度学习的图像识别算法。它和传统的目标检测方法不同,它将目标检测问题转化为为了一个回归问题^[8,9]。YOLO 算法将输入的图片分成了大小固定的小网格,并在每个小网格中预测围绕目标一周的边界框的位置和类别概率,这样它只需进行一次前向传播即可同时预测出所有目标的位置和类别。这使得 YOLO 具有实时性和快速性,能够在视频或图像中准确识别和定位多个目标物体。

YOLO 算法将图像划分为网络特征图,并使用目标的中心点所在的网格来生成目标的边框图。随着 YOLO 系列的不断发展,后续的版本如 YOLOv2^[10]、YOLOv3^[11]、YOLOv4^[12]、YOLOv5^[13]等都在不断优化精度和速度,以适应更多的应用场景^[8]。表 1 展示了 YOLO 关键版本及其特性。

表 1 YOLO 版本特性表

| 版本名 | 版本特性 |
|------------------------|---|
| YOLOv1 ^[7] | YOLO 算法的第一个版本,它首次提出了单次前向传播的思想,将目标检测问题转化为为了回归问题。 |
| YOLOv2 ^[10] | 在 YOLOv1 的基础上进行了小幅度的改进,引入了多尺度特征提取和卷积核大小变化的思想,还提高了检测的精度和速度。 |
| YOLOv4 ^[12] | 在前代的基础上进行了改良,引入了 CSPDarknet53 骨干网络、SPP 和 PAN 结构,同时提高了检测精度和速度。 |
| YOLOv8 ^[14] | 对比前代版本改动不多,但是大幅提升了模型的性能。 |
| YOLOv9 ^[15] | 最新最强的 YOLO 算法,引入了可编程梯度信息(PGI)和基于梯度路径规划的通用高效层聚合网络(GELAN)。 |

由图 1 可以看出,YOLOv8 在与近代 YOLO 的对比中,精度提升很多,速度也是提升明显。YOLOv8 各种指标全面超越现有对象检测与实例分割模型,它不仅有着 YOLOv5^[13]、YOLOv6^[16]、YOLOX^[17]等老模型的设计优点,而且在全面提升改良了 YOLOv5 模型结构的基础上同时保持了 YOLOv5 工程化简洁易用的优势。

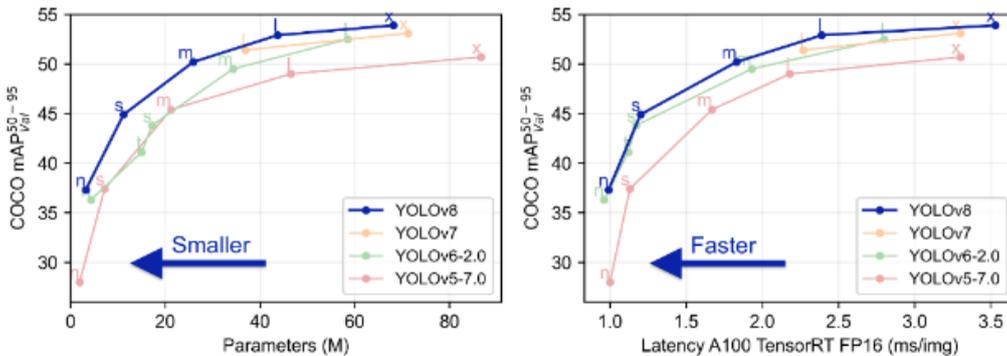


图 1 YOLOv8与近代YOLO的精度和速度对比^[14]

2.3 YOLOv8 模型结构

YOLOv8 模型相较于前代版本提升了性能与灵活性。YOLOv8 模型主要由四个部分组成,分别是:输入(input)、检测头(head)、主干网络(Backbone)、特征增强网络(neck)。YOLOv8 模型将输入的图片分为 $s \times s$ 个小网格,并让每个网格都只预测一个物体并产生 n 个边界框,所产生的边界框与每个网格相关。这样,YOLOv8 模型预测边界框的位置、大小和相关物体的概率,然后根据预测结果生成目标检测结果。YOLOv8 的网络结构如图 2 所示。

当前,基于 YOLO 的蛇类识别研究未见报道,本研究采用 YOLOv8 算法进行蛇类识别。

2.4 YOLOv8 算法原理

目标检测要完成三个任务:目标位置检测,目标大小检测和目标类别识别。于是,YOLO 算法按照如下步骤完成这 3 个任务。

(1) 网格分割。YOLO 算法将图片分割成大小相等的 $s \times s$ 个小网格,并识别中心落在网格之中的物体。如此算法并不需要特别大的网格框住整个物体,例如图 3 所示,一张 448×448 的图片按 3×3 被分割成 9 个小网格。

(2) 边界框表示。YOLO 算法预测出若干个边界框,预测出的边界框可表示为 5 个数据:识别出的物体的中心位置(x,y)、高(h)、宽(w)和预测的置信度,

以及物体类别。每个小网格中有多个边界框，但是只识别出一个物体。

(3) 数据处理。YOLO 算法输出的位置坐标是一个正实数，为了避免模型在大小不同物体上泛化能

力的差异过大，要对数据进行归一化。x 需要除以小网格的宽度，y 也需要除以小网格的高度；因为物体有可能大于小网格的大小，所以，h 要除以整个图片的宽度，同样的，w 要除以整个图片的高度。

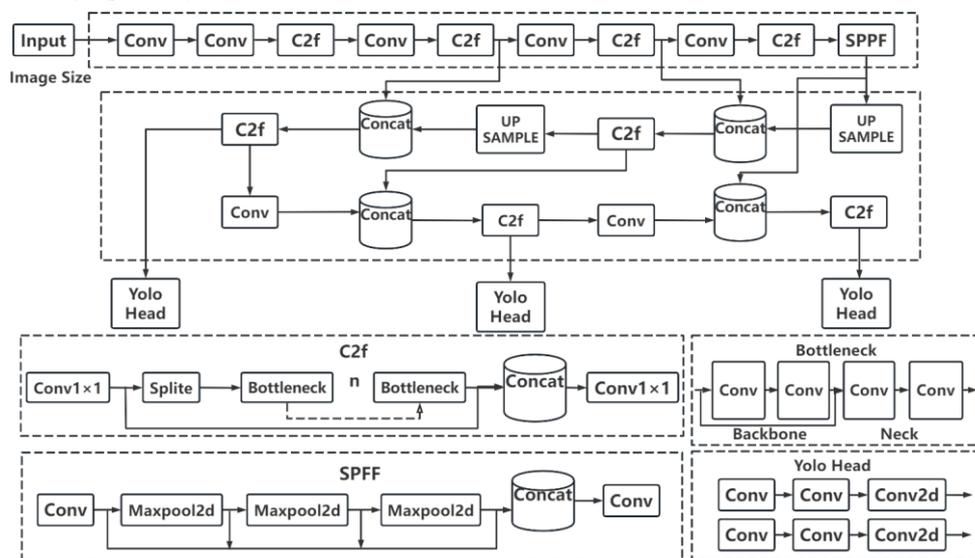


图 2 YOLOv8的网络结构

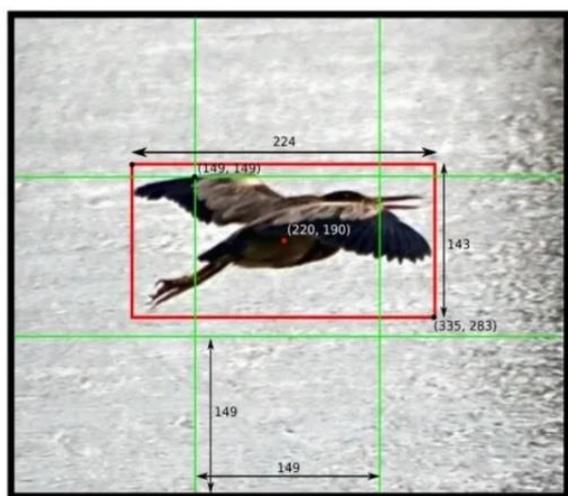


图 3 网格分割例图

(4) 真实值和置信度计算。对于 $s \times s$ 个小网格，每个小网格的边界框个数为 b ，分类器可以识别 c 种物体，则真实值总长度为 $s \times s \times (b \times 5 + c)$ 。用 IOU_{truth}^{pred} 表示预测的边界框和真实的物体位置的交并比，令 $Pr(Obj)$ 是一个小网格中有物体的概率，有物体时值为 1，没有物体时值为 0。于是，计算置信值 $C = Pr(Obj) * IOU_{truth}^{pred}$ 。

(5) 小网格选择。YOLO 算法运用非极大值抑制算法选择最符合要求的小网格。通俗地说，非极大值抑制算法就是保留最大置信值，而抑制其余置信值来预测有多少可能物体在当前小网格中。

(6) 物体识别。YOLO 算法将框选物体的边界框的置信值和该边界框所属的小网格的类别概率相乘，得到小网格内是该物体的概率，从而求出物体识别的概率以识别物体。

3 基于 YOLOv8 的蛇类识别系统

3.1 蛇类数据集

本研究可识别 15 种蛇类，蛇类名称中英文对照如表 2 所示。本研究使用 7000 张图片按照 7:2:1 的比例划分训练集、验证集和测试集，并以此来训练模型 YOLOv8。

表 2 所识别蛇类中英文名称对照如表

| 所识别蛇类英文名 | 所识别蛇类中文名 |
|-----------------------------------|----------|
| Mangrove cat snake | 红树林猫蛇 |
| Northern short-headed snake | 北部短头蛇 |
| Paradise tree snake | 天堂树蛇 |
| Philippine blunt-headed cat snake | 菲律宾钝头猫蛇 |
| Philippine bronze back tree snake | 菲律宾青铜背树蛇 |
| Philippine cat snake | 菲律宾猫蛇 |
| Philippine cobra | 菲律宾蛇 |
| Philippine pit viper | 菲律宾蝮蛇 |
| Philippine stripe-lipped | 菲律宾条纹蛇 |
| Philippine whip snake | 菲律宾鞭蛇 |
| red-tailed green rat snake | 红尾绿鼠蛇 |
| reddish rat snake | 红鼠蛇 |
| Reinhardt's Lined Snake | 莱因哈特条纹蛇 |
| reticulated python | 网纹蟒蛇 |
| smooth-scaled mountain rat snake | 光鳞山鼠蛇 |

所采集的蛇类图像中包含了标签、蛇类本体和背景板，本研究对蛇类图像进行了数据清洗、数据标注和数据增强预处理。

3.2 蛇类识别系统设计

(1) 系统整体设计

蛇类识别系统为用户提供了图片检测、摄像头检测、视频检测等功能，其系统结构如图 4 所示。

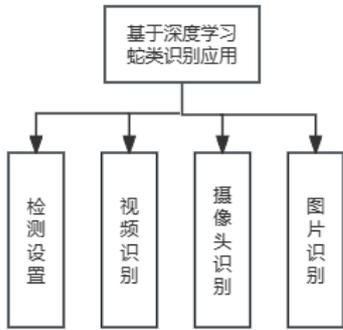


图 4 系统结构图

本系统主要围绕准确性、实时性、实用性和高效交互的设计目标实现基于 YOLOv8 的蛇类识别。

(2) 系统业务流程

本蛇类识别系统主要的功能是通过上传数据后对数据实现识别结果分析。系统功能模块总流程如图 5 所示。

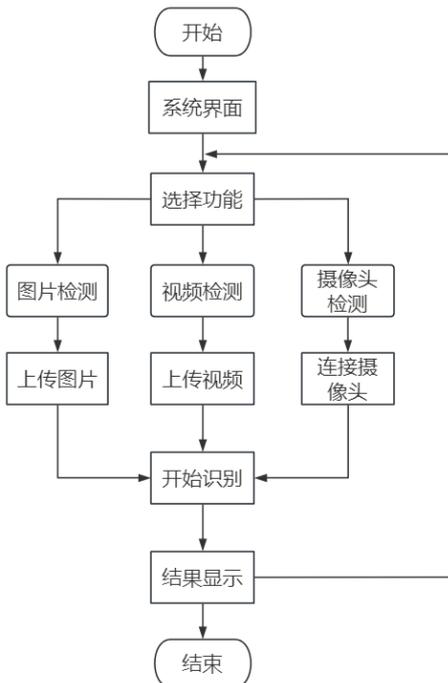


图 5 系统总流程

(3) 系统界面组成

根据系统总流程（如图 5），系统主要界面（如表 3 所示）设计为系统界面、蛇类识别界面、结果展示界面。

表 3 界面详细信息

| 界面 | 功能 |
|--------|-----------|
| 系统界面 | 选择功能、上传文件 |
| 蛇类识别界面 | 呈现所识别蛇类信息 |
| 结果展示界面 | 结果展示 |

4 蛇类识别系统实现

4.1 模型训练

本系统在如表 4 所示的实验平台进行模型训练，训练时所用的数据集来自网络数据库。

表 4 蛇类识别系统软硬件平台

| 名称 | 配置 |
|--------|--|
| CPU | Intel(R) Core(TM) i5-10300H CPU @ 2.50GHz 2.50 GHz |
| 内存 | 16GB |
| GPU | NVIDIA GeForce RTX2070 |
| CUDA版本 | CUDA 11.1 |
| 操作系统 | Win10(64位) |
| 环境 | Python3.6 |

4.2 蛇类识别系统界面展示

(1) 系统界面展示

本系统启动完成后进入系统界面（如图 6）。界面为用户提供图片检测，摄像头检测，视频检测等操作按钮，左边显示按钮，右边展示界面给出检测结果。

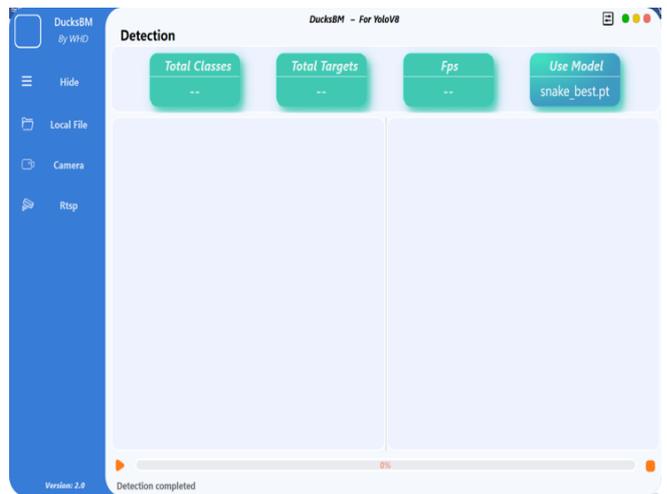


图 6 蛇类识别系统界面

图 7 呈现被识别蛇类的信息，其中 Total Classes 表示图片的数量，Total Targets 表示图片中识别出来

的蛇类个数，FPS 表示视频的帧率，Use Model 表示当前识别所用的模型。

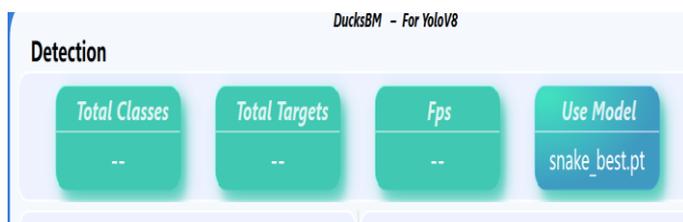


图 7 蛇类识别信息显示界面

由图 8 可知见，本系统可以在设置页面选择不同的蛇类识别模型，调整交并比(IOU)、置信度(Conf)，更改检测时间(Delay),和选择是否保存识别完成后的图片或视频的标签。

(2) 图片检测界面

本系统对图片进行检测需要经历输入与预处理、特征提取、网格划分与预测、非极大值抑制和检测结果输出 5 个环节。图片检测界面如图 9 所示。

(3) 摄像头检测界面

本系统的摄像头检测过程和图片检测过程类似。摄像头检测界面如图 10 所示。



图 9 图片检测界面

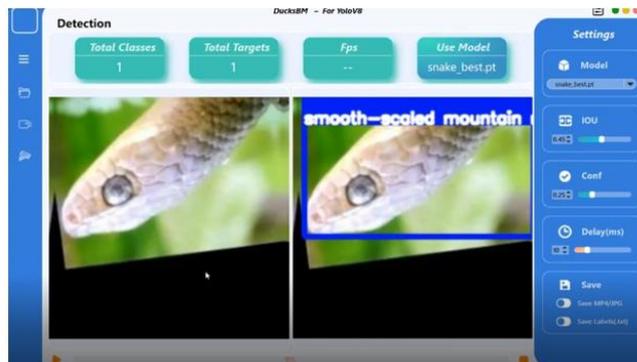


图 10 摄像头检测界面

(4) 视频检测界面

本系统的视频检测过程和图片检测过程类似。视频检测界面如图 11 所示。



图 11 视频检测界面



图 8 系统设置界面

6 结束语

本研究设计了一个蛇类识别系统，它集成了包括数据采集、预处理、建模以及识别结果展示等多种技术。本系统在技术上实现了准确快速的蛇类识别，它不仅能检测图片中蛇类，还能检测视频中或摄像头前的蛇类。本系统采用 YOLOv8 目标检测算法识别蛇类，运用 PythonGUI 图形界面 Pyside6 开发系统展示界面，

实现前后端稳定适配的连接, 为用户提供准确、实时、实用和高效的蛇类识别服务。

本系统虽在电脑端基本能够满足大多数用户对蛇类识别的需求, 但未来在此基础上可对模型轻量化以适应移动端运行的需要, 而且增加所识别蛇的习性和毒性信息展示, 此外改进系统的界面和用户交互设计, 以提高用户的交互体验, 另外还需优化深度学习模型, 加大训练集规模以进一步提高蛇类识别的准确性。

参考文献

- [1] 张乐, 陶明宝, 陈鸿平等. 常用蛇类药材鉴别研究进展[J]. 中国实验方剂学杂志, 2017, 23(4): 222-227.
- [2] 李章维, 胡安顺, 王晓飞. 基于视觉的目标检测方法综述[J]. 计算机工程与应用, 2020, 56(08).
- [3] 张皓洋. 基于 Inception-v3 模型的蛇类图像识别[J]. 电子技术与软件工程, 2019, 10, 58:61.
- [4] 付永钦. 基于深度学习的蛇类图像分类问题研究[D]. 学位论文, 浙江大学, 2019.
- [5] 鲁元平, 罗志聪. 基于 Squeezenet 模型的蛇类图像识别[J]. 电脑知识与技术. 2021, (7).
- [6] 周志斌, 罗志聪, 张展榜等. 基于迁移学习的中国蛇类识别研究[J]. 野生动物学报. 2022, 43(2): 436-443.
- [7] 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, Las Vegas, USA, 2016: 779-788.
- [8] 刘俊明, 孟卫华. 基于深度学习的单阶段目标检测算法研究综述[J]. 航空兵器, 2020, 27(03).
- [9] 邵延华, 张铎, 楚红雨等. 基于深度学习的 YOLO 目标检测综述[J]. 电子与信息学报, 2022, 44(10).
- [10] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, US: IEEE, 2017: 6517-6525.
- [11] REDMON J and FARHADI A. YOLOv3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H M. Yolov4: optimal speed and accuracy of object detection: arXiv: 2004.10934[R]. Ithaca, NY, US: Cornell University, 2020: 2004. 10934.
- [13] JOCHER G, STOKEN A, BOROVEC J, et al. Ultralytics/YOLOv5: V3.1-bug fixes and performance improvements [EB/OL].
- [14] Jocher G, Chaurasia A, Qiu J. Ultralytics YOLO (Version 8.0.0) [Computer software]. <https://github.com/ultralytics/ultralytics>, 2023.
- [15] Wang C, Yeh I, Mark Liao H. YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information[J]. arXiv:2402.13616v2, 2024.
- [16] LI C, LI L, JIANG H, et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications[J]. Ithaca: Cornell University Library, arXiv.org EISSN: 2331-8422 DOI: 10.48550/arxiv.2209.02976 arXiv.org, 2022-09.
- [17] GE Z, LIU S T, WANG F, et al. YOLOX: exceeding YOLO series in 2021[J]. arXiv:2107.08430[R]. Ithaca, NY, US: Cornell University, 2021:2107-08430.