

文章编号: 2095-2163(2022)08-0129-04

中图分类号: TP391.4

文献标志码: A

基于 YOLOv5 算法对斑马鱼幼鱼的检测研究

周福欢, 柴鑫雨

(上海理工大学 健康科学与工程学院, 上海 200093)

摘要: 斑马鱼幼鱼行为学分析常常是药物学研究及基因操作的直观表现, 在封闭、复杂环境下对幼鱼进行检测是研究其功能特性的基本步骤。由于幼鱼属于小目标, 本文通过去掉 YOLOv5 网络中的大尺度预测层和大、中尺度预测层, 得到了 YOLOv5m-sm 模型以及 YOLOv5m-s 模型; 由于没有公开的幼鱼数据集, 本文使用 DarkLabel 标注软件将幼鱼头部作为特征标记, 得到的 Zebdata 数据集, 并按 4:1 的比例分为训练集与验证集, 分别用来训练及验证模型; 为了测试算法对幼鱼的检测能力, 使用 160 张含有 23 条幼鱼的测试集对 YOLOv3m、YOLOv5s、YOLOv5m-s、YOLOv5m-sm 和 YOLOv5m 模型进行识别实验。实验结果表明, YOLOv5m-s 算法具有较高的识别准确度, 满足幼鱼目标检测要求。

关键词: 目标检测; YOLOv5; 斑马鱼幼鱼

A study on the detection of zebrafish larvae based on YOLOv5

ZHOU Fuhuan, CHAI Xinyu

(School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] Behavioral analysis of zebrafish larvae is often a visual representation of pharmacological studies and genetic manipulations. Therefore, the detection of zebrafish larvae in a closed and complex environment is a fundamental step in the study of their functional properties. Since the larva belong to small targets, YOLOv5m-sm model is obtained by removing the large prediction layer and YOLOv5m-s model is obtained by removing the large and medium prediction layer in YOLOv5m network. Meanwhile, since there is no publicly available larvae dataset, this paper uses DarkLabel software to label heads of the larva to obtain the Zebdata dataset that is divided into training and validation sets in a ratio of 4:1 to train and validate the model, respectively. In addition, 160 images that containing 23 larvae are used in this paper to test the detection accuracy of YOLOv3m, YOLOv5s, YOLOv5m-s, YOLOv5m-sm and YOLOv5m. The experimental results show that the YOLOv5m-s has high recognition accuracy and meets the requirements for larvae detection experiments.

[Key words] target detection; YOLOv5; zebrafish larvae

0 引言

斑马鱼幼鱼作为典型模型动物之一, 其视觉引导行为, 如: 视动反射、视运动反应、逃避反应等; 精神疾病引起的行为, 如: 药物成瘾和戒断、攻击性、恐惧和焦虑等得到了广泛的研究。视频检测是研究幼鱼行为的基础步骤。但是, 幼鱼体积小, 身体边缘透明, 与培养皿边缘对比度低, 运动模型处于近似睡眠或爆发状态, 使得识别难度大^[1]。传统的目标检测算法主要依靠目标物体与背景之间的差异来进行识别, 比如: 物体的灰度、颜色、光流、边缘、纹理和梯度等特征。但幼鱼姿态变化、不连续运动、遮挡、光照强度等环境的干扰使得检测精度低。近年来, 由于卷积神经网络对图像的强大表征能力, 也在一定程度上推动了目标检测算法的较为快速发展。

目前, 基于深度学习的检测算法一般可分为 2 类。一类是基于候选框的两阶段检测算法, 具有较

高的定位和识别精度, 如 R-CNN、Fast R-CNN、Faster R-CNN、Mask R-CNN; 另一类是基于回归的单阶段检测算法, 具有较高的检测速度, 如 YOLO 系列算法和 SSD 算法。其中, YOLO 算法已经成为多目标检测领域的研究热点, 现已广泛应用于各个行业^[2]。YOLO 可以将候选框定位、特征提取和物体分类整合到一个骨干网络中, 实现端到端的检测。以此为基点, YOLOv2、YOLOv3、YOLOv4 和 YOLOv5 也紧随其后相继问世, 其中 YOLOv5 具有轻量级架构, 检测精度上与 YOLOv4 相当, 但速度优于 YOLOv4。基于此, 本文将使用基于 YOLOv5 网络对斑马鱼幼鱼进行检测。

1 YOLOv5 网络模型介绍

YOLOv5 算法根据网络宽度与深度分为 4 种模型, 依次为 YOLOv5s、YOLOv5m、YOLOv5x 和 YOLOv5l, 精度逐渐提高, 但检测速度逐渐降低。本

作者简介: 周福欢(1996-), 女, 硕士研究生, 主要研究方向: 机器视觉; 柴鑫雨(1999-), 女, 硕士研究生, 主要研究方向: 机器视觉。

收稿日期: 2022-01-18

哈尔滨工业大学主办 ◆ 专题设计与应用

文选取 YOLOv5m 检测斑马鱼幼鱼,在有限的设备条件下获取最好的检测精度与检测速度。

YOLOv5 网络结构可分为 4 个部分,即:输入 (Input)、骨干网络 (Backbone)、颈部 (Neck) 和预测层 (Prediction),如图 1 所示。输入端对图片进行预处理,主要有自适应图片缩放、自适应锚框计算,并使用 Mosaic 来扩充数据集;骨干网络对输入图像提取特征信息;颈部的作用是融合从骨干网部分提取的特征信息,发送到预测层部分;预测层由 3 个尺度的特征图组成,分别用于检测小、中、大目标。例如,

当输入分辨率为 608×608 的图像时,将检测到分辨率超过 8×8 的小目标、分辨率超过 16×16 的中等目标,以及分辨率超过 32×32 的大目标。故使用 K-means++ 聚类算法获取数据集的锚框 (Anchor boxes),得出幼鱼的头部矩形框大小在 $8 \sim 20$ 像素之间。因此,本文对 YOLOv5 删除大尺度的预测层部分和其相应的颈部组件,保留小、中尺度预测层部分,得到 YOLOv5m-sm 模型,去掉大、中尺度预测层部分,保留小尺度预测层,得到 YOLOv5m-s 模型。

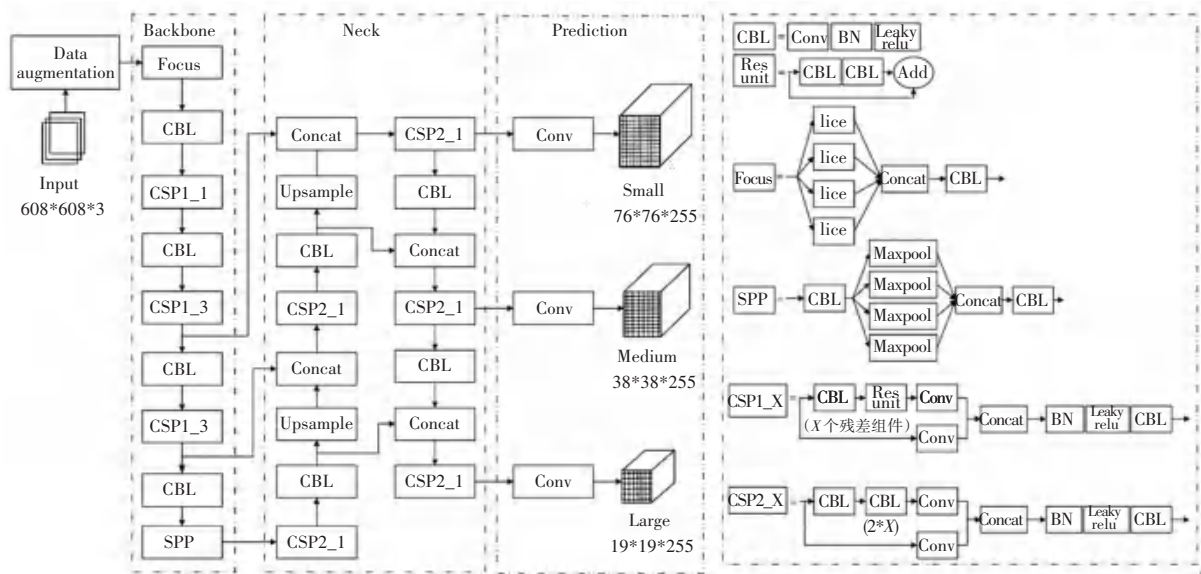


图 1 YOLOv5 结构图

Fig. 1 The structure of YOLOv5 model

2 实验

2.1 实验平台

本文算法的实验平台配置,见表 1。

表 1 实验平台配置

Tab. 1 Experimental platform configuration

名称	具体参数
处理器	Intel Core i7-8550u CPU 1.8 GHz
显卡	NVIDIA CUDA 10.2
内存	8 GB
操作系统	Windows 10 Professional

2.2 数据集

将幼鱼放在培养皿中,培养皿底部放置一个 LED 光源,手机以一定高度在室内光线下垂直拍摄,具体信息见表 2。拍摄的图像含有杂质、排泄物、波纹等来提高数据集的多样性。

本文使用 DarkLabel 标注软件以 1 fps 将视频进行分帧,选取幼鱼眼部进行标记,由此就得到了

Zebradata 数据集。每张图像标记后得到一个.txt 文件,主要包括目标类别、识别难度、矩形框的左上角与右下角坐标信息。

表 2 幼鱼数据集

Tab. 2 The dataset of larvae

年龄 (dpf)	数量	帧数	尺寸 (pixel)
5	3	1 057	544 * 544
		933	720 * 720
18	1	972	720 * 720
	3	305	720 * 720
22	1	504	544 * 544
	2	727	544 * 544
	4	218	720 * 720
24	4	835	720 * 720

为了增强数据多样性,提高模型训练时的泛化能力,使用了数据扩充方法,如: Mosaic、Mixup 和 Cutout。其中, Mosaic 是随机选取 4 张图片,按照随机缩放、随机裁剪和随机排布的方式进行拼接。

Mixup 是将多张图像像素融合形成一张新的图像。Cutout 是利用一个任意大小的矩形框随机遮挡原图像上的任意区域, 模拟目标遮挡情况。

2.3 实验流程

基于 YOLOv5 的斑马鱼幼鱼多目标检测过程如

图 2 所示。这里, 将 ZebraData 数据集的图像大小归一化为 640×640 , 批量大小设为 2, 初始学习率设为 0.003 2, 训练 50 世代 (epochs), 可使损失函数收敛。训练好的模型可对幼鱼自动识别与定位, 显示出目标类别与置信度。

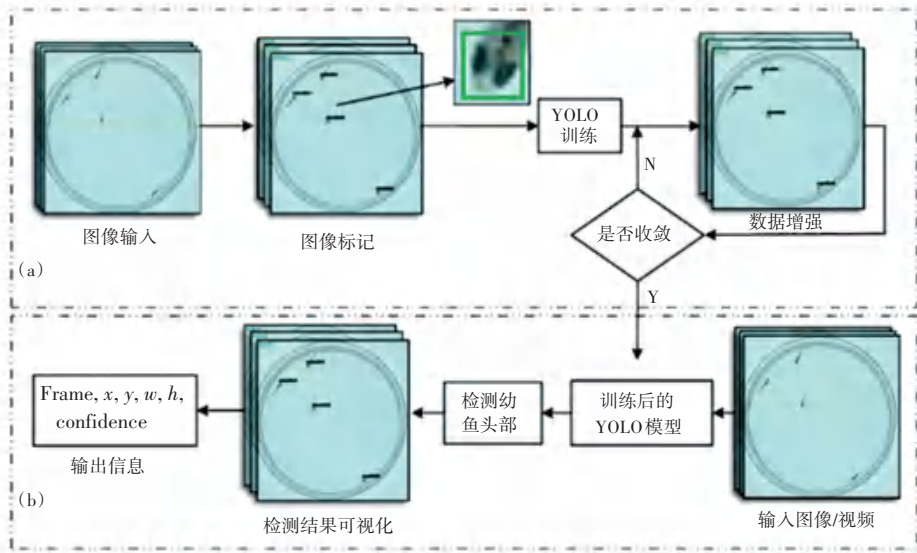


图 2 基于 YOLOv5 算法的幼鱼多目标检测过程

Fig. 2 Multi-target detection process for zebrafish larvae based on YOLOv5

2.4 评价准则

为了评估 YOLO 模型对幼鱼的检测能力, 本文采用准确率 (Precision)、召回率 (Recall) 和平均精度 (AP) 指标对模型进行评估。对此可做研究分述如下。

(1) 准确率与召回率。计算公式具体如下:

$$Precision = \frac{T_p}{T_p + F_p} \quad (1)$$

$$Recall = \frac{T_p}{T_p + F_n} \quad (2)$$

其中, T_p 表示目标被正确预测出来的数量; F_p 表示目标被误测出来的数量, 即误检; F_n 表示目标未被预测出来的数量, 即漏检。

(2) 平均精度 (Average Precision, AP): 用于评估一个类别训练的模型, 主要是通过计算模型的 P-R 曲线下的面积, 此时需用到的数学公式为:

$$AP = \int P(R) d(R) \quad (3)$$

其中, P 为精确率, R 为召回率。

3 实验结果与分析

使用 160 张含有 23 条幼鱼的图像作为测试集,

批量大小为 4, 置信度阈值为 0.5, 交并比 (IOU) 阈值为 0.9, 图像输入大小归一化为 736×736 。为了体现本文算法性能的优越性, 对 YOLOv3m、YOLOv5s、YOLOv5m-s、YOLOv5m-sm 和 YOLOv5m 模型进行精度测试实验。实验结果见表 3。

从表 3 中可以看出, YOLOv3 网络检测精度不及 YOLOv5 网络, 只含有小尺度预测层的 YOLOv5m-s 模型性能表现最好, 相对于 YOLOv5m, 准确度提高 19.71%, 回收率提高了 10.34%, 平均精度上提高 13.29%, 推理速度 (Inference) 提高 13.92%。当幼鱼处在培养皿边缘时, 与背景对比度很低的情况下, YOLOv5m-s 模型也可以很好地检测到幼鱼。

表 3 检测结果

Tab. 3 The experimental detection results

模型	准确率	回收率	平均精度	推理速度/ms
YOLOv3m	0.699	0.699	0.745	259.5/3.21
YOLOv5s	0.686	0.676	0.710	71.4/1.40
YOLOv5m-s	0.923	0.886	0.912	160.2/1.30
YOLOv5m-sm	0.840	0.804	0.815	174.2/1.40
YOLOv5m	0.771	0.803	0.805	186.1/1.40