

DOI:10.13718/j.cnki.xsxb.2023.08.011

# 基于改进 YOLOv5 的鸽子蛋壳破损检测<sup>①</sup>

杨航, 何皓明, 李滕科,  
王嘉雯, 吴霆, 钟乐, 邹娟, 杨灵

仲恺农业工程学院 信息科学与技术学院/广东省食品安全溯源与控制工程技术研究中心, 广州 510225

**摘要:** 破损鸽蛋检测是鸽蛋分拣过程中最重要的步骤之一。为解决破损鸽蛋实时检测问题, 提出一种改进 YOLOv5 的破损鸽蛋检测方法。首先, 为解决鸽蛋细微裂纹检测难的问题, 采用 Kmeans 聚类算法计算细微裂纹锚框适应度, 在 head 模块增加 Anchor 值为 $[7, 9, 13, 8, 9, 12]$ 的检测层, 从而提高细微裂纹图像的特征提取能力; 其次, 为解决裂纹特征权重较低的问题, 采用融合反向最大池化层的注意力机制, 将色值权重提高, 从而增强裂纹特征的权重; 再次, 为解决检测模型负向梯度消失问题, 采用 Hardswish 激活函数的全连接层替换 ReLU 激活函数的全连接层, 解决输入为负时存在的神经元坏死问题, 确保神经网络误差正常反馈。最后, 基于改进 YOLOv5 方法建立鸽蛋裂纹识别模型, 采用 7:2:1 交叉验证方法对模型进行验证。经对比验证, 改进 YOLOv5 模型的检测效果有明显提升, 其准确率、召回率和平均精度分别达到 98%, 97.3% 和 98.3%, 比原始 YOLOv5 模型分别提高了 14.8%, 6.2% 和 3.4%, 检测速度大约为 111 帧/s, 且模型大小仅为 4.6 Mb, 为鸽蛋破损检测提供了一种新方法。

**关键词:** 鸽子蛋; YOLOv5; 裂纹检测; Hardswish; 最大池化层

中图分类号: TP181

文献标志码: A

文章编号: 1000-5471(2023)08-0092-11

## Detection of Pigeon Eggshell Breakage Based on Improved YOLOv5

YANG Hang, HE Haoming, LI Tengke, WANG Jiawen,  
WU Ting, ZHONG Le, ZOU Juan, YANG Ling

School of Information Science and Technology, Zhongkai University of Agriculture and Engineering/

Guangdong Provincial Food Safety Traceability and Control Engineering Technology Research Center, Guangzhou 510225, China

**Abstract:** Pigeon egg broken detection is one of the most important steps in the pigeon egg sorting process. In order to solve the problem of real-time detection of damaged pigeon eggs, this paper proposed a method to improve YOLOv5's detection of damaged pigeon eggs. First, in order to solve the difficulty of pigeon egg fine crack detection, the Kmeans clustering algorithm was used to calculate the fitness of the fine crack anchor frame, and the detection layer with the anchor value of  $[7, 9, 13, 8, 9, 12]$  was added to the head module to improve the feature extraction ability of the fine crack image; Secondly, in order to solve the

① 收稿日期: 2022-12-27

基金项目: 广东省重点领域研发计划项目(2021B0202030001, 2019B020215001); 国家自然科学基金项目(31872606); 广州市重点研发计划项目(202103000067, 201803020033, 202002030154); 广东省教育厅乡村振兴重点领域专项(2020ZDZX1060); 广东省自然科学基金(2018A0303130034, 2020A1515010834); 广东省教育厅创新强校特色创新项目(2018KTSCX096); 广东省省级科技计划项目(2017A020225042); 广东省现代农业产业技术体系创新团队建设专项(2019KJ141, 2020KJ138)。

作者简介: 杨航, 硕士研究生, 主要从事人工智能研究。

通信作者: 邹娟, 讲师, 硕士研究生导师。

problem of low weight of crack characteristics, the attention mechanism of fusing the reverse maximum pool layer was adopted to increase the weights of crack color value, thus enhancing the weight of crack characteristics; Thirdly, in order to solve the problem of the disappearance of the negative gradient of the detection model, the full connection layer of the Hardswish activation function was used to replace the full connection layer of the ReLU activation function, to solve the problem of neuron necrosis when the input is negative, and to ensure the normal feedback of the neural network error. Finally, the pigeon egg crack identification model was established based on the improved YOLOv5 method, and the 7:2:1 cross validation method was used to verify the model. Through comparison and verification, the detection effect of the improved model of YOLOv5 in this paper has been significantly improved, with its accuracy, recall and average accuracy reaching 98%, 97.3% and 98.3% respectively, which was 14.8%, 6.2% and 3.4% higher than the YOLOv5. The detection speed was about 111 frames/s, and the model size was only 4.6Mb. This provides a new method for detecting broken pigeon eggs.

**Key words:** pigeon egg; YOLOv5; crack detection; Maxpool

在禽蛋经营中, 破损禽蛋分拣是一个非常重要的环节. 为解决传统人工分拣准确度低、漏检率高、速度慢的问题, 提高自动分拣的水平, 研究人员一直在研究采用先进技术来实现禽蛋破损自动检测. 在过去的 20 年里, 有不少专家学者将传统的机器视觉<sup>[1-2]</sup>和音频分析<sup>[3-4]</sup>等各种技术手段应用在禽蛋破损检测上. 贺静等<sup>[5]</sup>提出的基于 DSP 实时图像采集处理的鸡蛋破损检测方法, 将 DSP 实时采集的鸡蛋图片灰度化处理后, 应用外接矩形法描绘鸡蛋近似椭圆投影的外切矩形, 利用改进型标记法判断鸡蛋是否破损. 但该方法有一定的局限性, 检测时会受到蛋壳斑点的影响, 当斑点面积大于  $15.04 \text{ mm}^2$  时, 就会错检. 王树才等<sup>[6]</sup>根据声波在不同介质中的不同特性, 利用敲击声音信号进行研究, 发现正常蛋、破损蛋、钢壳蛋和尖嘴蛋在衰减时间、共振峰频率、最大频差 3 项指标上存在显著差异, 据此实现蛋的破损检测. 但该方法受到噪声和敲击点的位置影响较大, 对检测精度有一定的影响.

从相关研究文献<sup>[1-6]</sup>可以看出, 传统机器学习算法应用在禽蛋破损检测时, 鲁棒性比较差, 容易受到多因素的干扰, 漏检和错检问题比较突出, 很难满足大规模分拣的需要. 近年来, 计算机视觉技术迅猛发展, 目标检测算法不断改进和创新, 被广泛运用在各个领域中<sup>[7-8]</sup>. 在蛋破损检测领域, 研究人员也不断尝试将各种目标检测算法应用在研究中, 极大地提高了蛋壳破损检测的精度和速度. 如涂伟沪等<sup>[9]</sup>利用蝗虫算法 (grasshopper optimization algorithm, GOA) 改进 canny 算子, 提高了鸡蛋线型、网状裂纹的检测精度, 对鸡蛋的线型和网状裂纹的漏检率分别降低了 20% 以上和 60% 以上; 赵祚喜等<sup>[10]</sup>用 YOLOv4 进行了鸡蛋破损的在线检测研究, 破损蛋的识别率达 85% 左右.

相比于鸡、鸭、鹅等禽蛋, 鸽子蛋具有体积更小、蛋壳更易碎的特点. 据汤青萍等<sup>[11]</sup>人研究, 鸽子蛋壳平均厚度为 0.23 mm, 蛋壳强度 1.19 kg/cm, 乳突层稀疏、形状不规则且直径较大, 栅栏层断面凹凸不平, 晶体层晶体排列松散、晶格棱角模糊. 基于此, 本文实验采用目前综合性能最优的目标检测算法 YOLOv5 来研究鸽子蛋壳破损检测. 以自行设计的设备分别对裂纹鸽蛋和完好鸽蛋的图像进行采集, 对数据通过 Mosaic 数据增强<sup>[12]</sup>、自适应锚框计算、自适应图片缩放等进行预处理. 在实验过程中, 出现了模型对部分裂纹鸽蛋识别准确率不高的问题, 原因在于细微裂纹在深层网络中出现特征丢失、气室特征权重过高的问题, 干扰了裂纹特征的提取. 为此, 本研究对 YOLOv5 模型进行了相应的改进, 增加一个检测层以提高对细微裂纹特征的提取能力, 融合改进的 SE 注意力机制 (squeeze-and-excite)<sup>[13]</sup>以提高裂纹特征的权重. 本文采用 YOLOv5<sup>[14]</sup>, Faster RCNN<sup>[15]</sup>, YOLOv3-SPP<sup>[16]</sup>进行对比, 结果显示, 本文改进的 YOLOv5 模型在鸽蛋破损检测中综合优势明显, 为鸽蛋的破损检测提供了新的解决方案.

## 1 模型选择

### 1.1 YOLOv5 算法描述

YOLOv5 由 4 个模块组成: 输入端 (Input)、主干网络 (Backbone)、颈部网络 (Neck) 和输出端 (Head), 其网络结构见图 1.

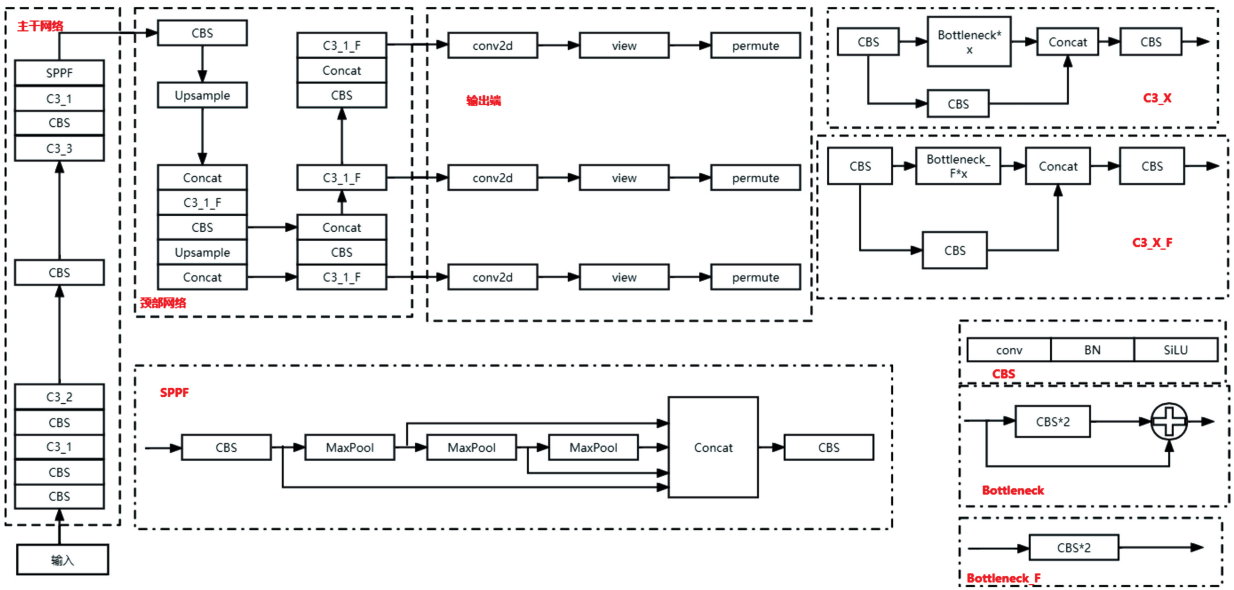


图 1 YOLOv5 网络结构

### 1.1.1 输入端

主要功能是数据输入与数据增强,包括 Mosaic 数据增强、自适应锚框计算、自适应图片缩放等.其中, Mosaic 数据增强通过对输入图像进行随机缩放、裁剪、拼接、排布等方式增加小目标的样本数量,从而提高模型的检测精度.

### 1.1.2 主干网络

主要包括 CBS、带 shortcut 的 C3 和 SPPF 3 种模块(图 1). CBS 的结构为 Conv+BN+SiLU,主要用于卷积操作; C3 为 CSP 架构<sup>[17]</sup>,包含了 3 个标准卷积层和多个 bottleneck 模块,这个部分的 C3 模块包含的位置信息、细节信息较多,但语义信息较少,主要用于特征提取; SPPF 模块采用了多个小尺寸池化核来代替单个的大尺寸池化核,在保留了单个大尺寸池化核的多感受野特征融合功能的情况下,进一步提高了运行速率.

### 1.1.3 颈部网络

位于主干网络和输出端之间,主要由 CBS、Upsample、Concat 和不带 shortcut 的 C3 组成.颈部网络采用 FPN+PAN 的结构,FPN 结构<sup>[18]</sup>能使所有尺度下的特征都有丰富的语义信息; PAN 结构<sup>[19]</sup>则主要用于加强定位信息. Neck 的 CBS 主要作用是通过下采样来抽取高层次语义信息. C3 则主要用于纹理特征的提取,包含的位置信息、细节信息较少,语义信息较多,将提取出的特征进行混合组合后,传递到预测层,增强网络特征融合的能力.

### 1.1.4 输出端

Head 模块利用之前获取的特征作出预测,采用 GIoU\_Loss 作为损失函数,同时通过非极大值抑制 NMS 来筛选目标框<sup>[20]</sup>.

因网络深度和维度不同, YOLOv5 模型主要有 YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x 等版本.其中, YOLOv5n 是支持移动端的最小化模型,训练出来的最终模型的大小是 YOLOv5s 的 1/4 左右,只有 3.4 MB,具有轻量化、精度高、速度快的特点.为满足轻量化要求,本研究实验选择了适合搭载在移动设备上、更轻量的 YOLOv5n 模型.

## 1.2 算法优化

针对鸡蛋破损检测过程中,裂纹目标细微、裂纹特征模糊、裂纹特征提取能力差等特点,本研究提出了两点改进.一是为提高细微裂纹检测精度,采用 Kmeans 聚类算法计算细微裂纹锚框适应度,并在 head 模块增加检测层,从而检测到更小的裂纹;二是为提高对鸡蛋纹理特征的提取和融合能力,用融合了改进 SE 注意力机制的 C3 模块代替原有的 C3 模块,提高裂纹特征的权重,增强对鸡蛋纹理特征的提取和融合

能力.

### 1.2.1 增加检测层

YOLOv5 的 3 个检测层对应 3 组初始化 Anchor 值[10, 13, 16, 30, 33, 23]、[30, 61, 62, 45, 59, 119]、[116, 90, 156, 198, 373, 326], 最大能检测  $32 \times 32$  的目标, 最小能检测  $8 \times 8$  的目标. 由于大部分鸽蛋的裂纹比较细微, 图像分辨率低、特征表达能力弱, 能提取的特征比较少, 更容易出现错检问题.

为此, 本研究采用 Kmeans 聚类算法计算细微裂纹锚框适应度, 迭代次数为 1 000, 获得自适应锚框 [7, 9, 13, 8, 9, 12]. 即为了检测和提取到更细微的裂纹特征, 在 head 模块中增加 Anchor 值为 [7, 9, 13, 8, 9, 12] 的检测层. 计算过程为:

$$d(box, anchor) = 1 - IOU(box, anchor) \quad (1)$$

$$IOU(box, anchor) = \frac{box \cap anchor}{box \cup anchor} \quad (2)$$

依此改进后, 主干网络的第 2 层就会开始增强特征, 并在 17 层后继续对特征图进行上采样等处理, 扩大特征图. 在第 20 层将获取到的大小为  $160 \times 160$  的特征图与主干网络中第 2 层特征图进行 concat 融合, 以此获取更大的特征图, 从而能够提取到  $4 \times 4$  大小的特征, 实现对比较细微的裂纹的特征提取.

### 1.2.2 添加注意力机制 SE 模块

在本鸽子蛋破损检测任务中, 细微裂纹的特征信息容易在深层网络中丢失. 同时, 在暗室照蛋的条件下, 气室部位会形成光晕, 给裂纹特征信息带来极大的干扰, 需要更加注重纹理特征的提取. SE 注意力机制<sup>[13]</sup>能通过建模来确定各个特征通道的重要程度, 然后针对不同的任务增强或者抑制不同的通道, 增强对裂纹特征信息的学习, 增加裂纹特征的权重, 可以让网络更加关注待检测目标, 减少细微裂纹特征信息在深层网络中丢失的现象, 从而提高检测效果. 为此, 本实验将 SE 注意力机制融合到 C3 模块中, 以增强裂纹特征信息的提取.

SE 注意力机制结构主要包含 Squeeze 和 Excitation 两部分.

Squeeze 通过全局平均池化层(global pooling)将全局空间信息压缩到信道描述符中, 将  $H \times W$  个像素压缩成一组  $1 \times 1 \times 3$  的实数, 计算公式为:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (3)$$

然后通过 Excitation 的激励操作, 将  $1 \times 1 \times 3$  的实数降维为  $1 \times 1 \times (3/r)$ , 然后再升维到  $1 \times 1 \times 3$ , 最后经过 sigmoid 激活函数的全连接层的归一化操作, 变成一组 0-1 的实数, 计算公式为:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (4)$$

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \max(0, W_1 z)) \quad (5)$$

其中  $\delta$  为 ReLU 激活函数.

这组 0-1 的实数对应了每个通道的重要性, 0 为最不重要, 1 为最重要. 最后再对应到特征图的每一个像素值, 得到的最终输出为:

$$F = (H \times W \times C) \times (1 \times 1 \times C) \quad (6)$$

考虑到大部分的裂纹在照蛋环境下呈现黑色纹路, 在 Excitation 的激励操作的降维中会出现得到的实数基本相同, 而裂纹部分的 RGB 图色值大多数接近  $\{0, 0, 0\}$ , 使得这部分的特征在归一化之后获得的权重很低, 出现这部分特征在深层网络中丢失的现象. 为此, 本实验对 SE 注意力机制的全局平均池化层进行修改, 以反向最大池化层的方式增强裂纹特征的权重. 改进后的 Squeeze 操作为:

$$z_c = F_{sq}(1 - u_c) = \max(\sum_{i=1}^H \sum_{j=1}^W 1 - u_c(i, j)) \quad (7)$$

然而, 这一改动会在降维操作中出现输入为负的现象. 如果继续使用 ReLU 激活函数的全连接层, 会出现小于零的输入存在时导致的神经元坏死问题(图 2).

于是, 本研究采用 swish 激活函数的全连接层, 这样即便在降维操作中输入为负也不会出现输出为 0 的现象, 可以较好地解决融合了反向的最大池化层后出现的负向梯度消失问题, 避免部分神经元坏死, 确保神经网络误差正常反馈(图 3).

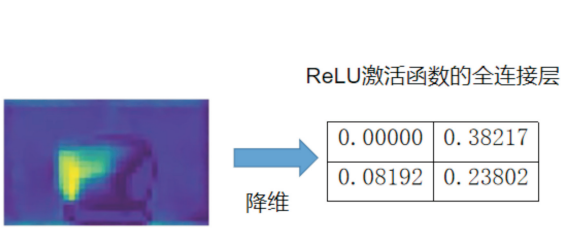


图 2 激活函数 ReLU 的全连接层降维结果

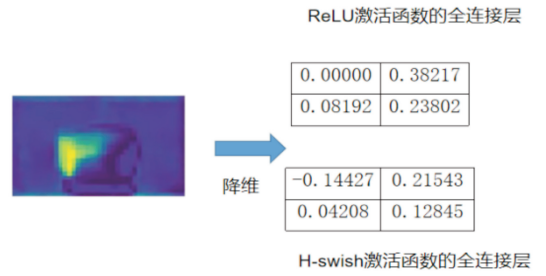


图 3 两种激活函数的降维结果对比

与 ReLU 激活函数相比, swish 激活函数<sup>[21-22]</sup>是一个平滑的、非单调的函数, 它和它的导数定义为:

$$f(x) = x \cdot \beta \cdot \sigma(x) \tag{8}$$

$$f'(x) = \sigma(\beta x) + \beta x \cdot \sigma(\beta x)(1 - \sigma(\beta x)) \tag{9}$$

其中:  $\sigma(x) = \text{sigmoid}(x)$ .

从中可知, 在  $x > 0$  时, swish 激活函数不会出现梯度消失的问题; 当  $x < 0$  时, 也不会出现 ReLU 激活函数导致的神经元坏死的问题. 在 SE 注意力机制的全连接层中, 采用 swish 激活函数的全连接层便不会出现得到的实数中出现 0 的现象, 由此可避免这部分特征在深层网络中丢失. 同时, swish 激活函数的导数也不像 ReLU 激活函数的导数那样一成不变, 而是处处可导、连续光滑, 能显著提高神经网络的准确性<sup>[23]</sup>, 让模型获得更高的精度.

虽然 swish 激活函数能提高精度, 但在移动设备和嵌入式设备上它所需的计算成本要比 ReLU 高得多. 在相同数据集下, 对分别使用 ReLU、swish 和 Hardswish 激活函数时模型的推理速度进行对比时发现, 使用 ReLU 激活函数时的单帧推理时间为 10.3ms, 使用 swish 激活函数时的单帧推理时间为 10.9 ms, 使用 Hardswish 激活函数时的单帧推理时间为 10.5ms. 使用 ReLU 激活函数和 Hardswish 激活函数时, 模型的推理速度差别不大, 降低了不到 2%; 而使用 swish 激活函数时, 模型的推理速度却降低了约 6%. 因此, 本文采用 Hardswish 激活函数代替 ReLU 激活函数<sup>[24]</sup>, 其定义为:

$$\text{Hardswish}(x) = x \cdot \text{HardSigmoid}(x) = x \cdot \frac{\text{ReLU6}(x + 3)}{6} \tag{10}$$

ReLU6(x)几乎可以在所有的软件和硬件框架上使用<sup>[24]</sup>, 在量化过程中消除了近似 Sigmoid 函数的不同实现而导致的潜在数值精度损失. 此外, Hardswish 激活函数可以实现分段功能, 以减少内存访问次数, 能使网络的推理速度加快, 对量化过程更加友好<sup>[25-26]</sup>.

ReLU, swish, Hardswish 3 种激活函数及其导数见图 4.

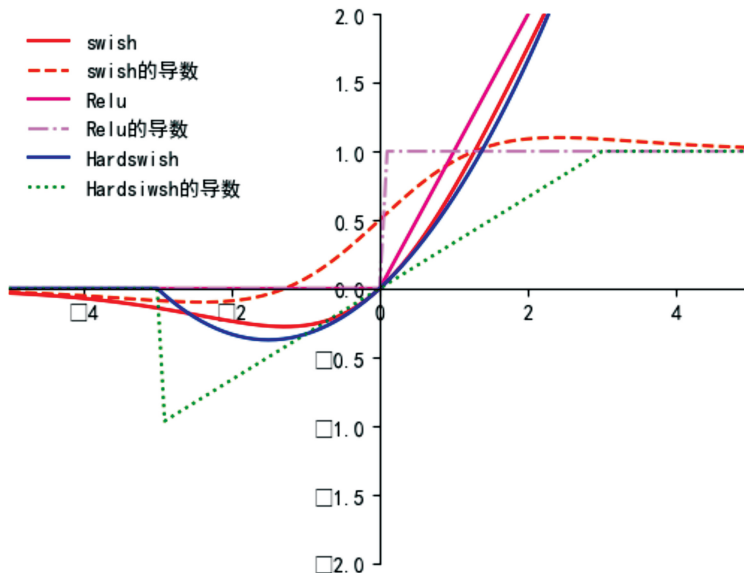


图 4 ReLU, swish, Hardswish 激活函数及其导数

改进后的 Excitation 的激励操作为:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \eta(W_1 z)) \tag{11}$$

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 2(W_1 z) * \max(0, \min(1, (0.2(W_1 z) + 0.5)))) \tag{12}$$

其中  $\eta$  为 Hardswish 激活函数。

改进后的 SE 注意力机制的结构见图 5。

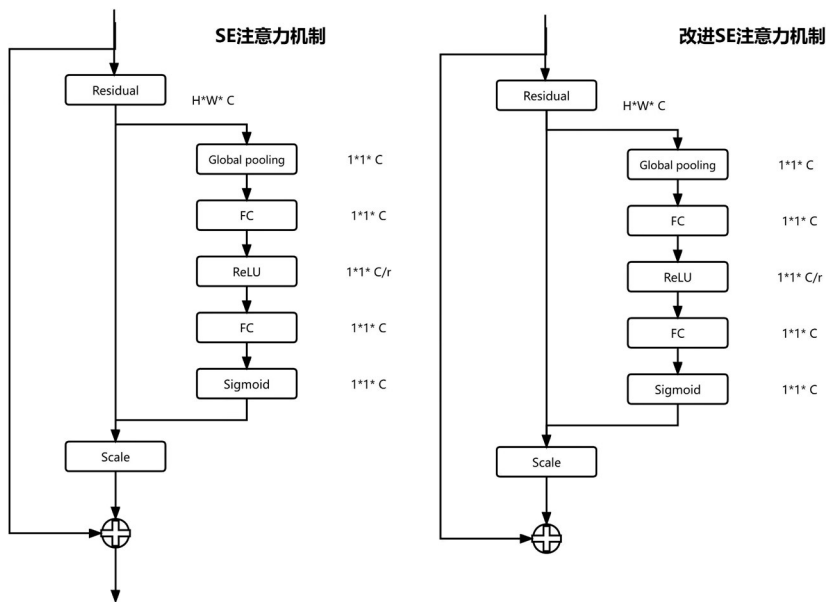


图 5 SE 注意力机制以及改进后的 SE 注意力机制

改进后 C3 结构见图 6。

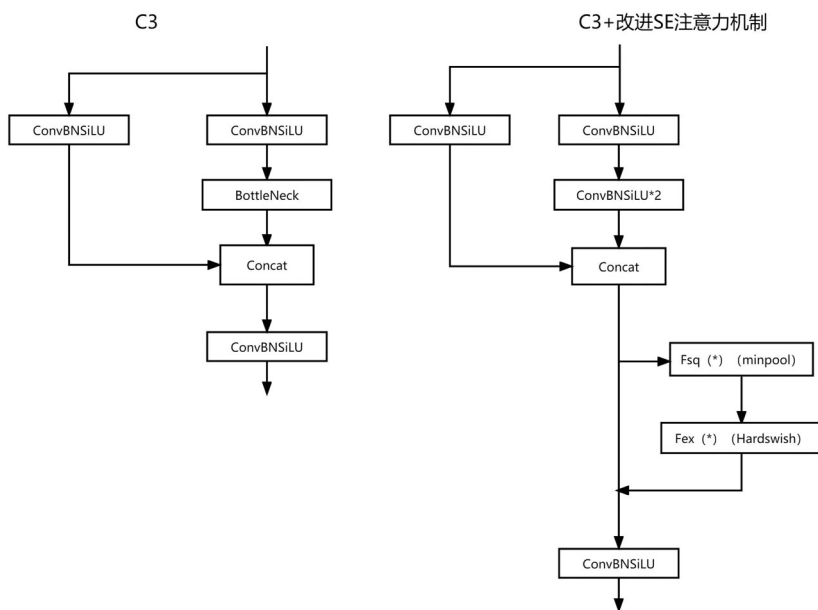


图 6 改进后 C3 的结构

## 2 实验过程

### 2.1 实验设备

本研究实验采用自主设计的数据采集设备, 旨在创造一个尽可能去除外部干扰的检测环境, 用来采集鸽蛋图像和实现对鸽蛋的快速检测。装置的主体部分由可控制传送带、暗室、摄像机、开发板组成。传送带

采用步进电机驱动,使用 STM32F030R8 芯片控制传送带的转动速度和识别单个鸽子蛋的时间;开发板使用的操作系统为 Ubuntu 18.04,检测 GPU 为 GTX760,显存容量 2048M,测试环境为 Python3.8,测试框架为 Pytorch1.9.0,使用 CUDA 版本 10.2.89 作为框架计算。

鸽子蛋通过传送带被连续不断送入暗室,每个鸽子蛋在固定时间内绕轴心旋转 360°,利用摄像头通过照蛋法采集到每个鸽子蛋的图像,每个鸽子蛋获得蛋身不同侧面的图片若干张,利用这些图片建立数据集并训练模型,最后由开发板搭载训练好的模型对鸽子蛋进行实时检测。实验装置及示意图见图 7。

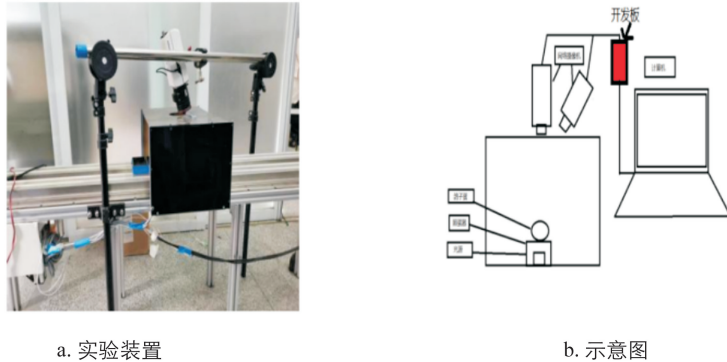


图 7 实验装置及示意图

## 2.2 数据采集和数据集建立

本研究的实验材料为 300 枚完好无损鸽子蛋,从鸽子养殖场和菜市场采购获得。先将 300 枚鸽子蛋通过图 7 实验装置拍摄获得图像,并将这些图像编组为“good”。然后用不同工具、不同力度在 300 枚鸽子蛋的不同部位敲出裂缝,尽量使裂缝的大小、形状、位置不一,并通过图 7 实验装置拍摄获得图像,编组为“crack”。

将“good”组和“crack”组图片建立数据集,并对“crack”组图像的裂纹通过图像翻转、旋转及对裂纹复制、粘贴等方式进行数据增强,对数据集进行扩容,得到 2179 张图片组成的数据集,其中将 1 620 张作为训练集,559 张作为测试集。

## 2.3 训练环境以及模型训练

在网络模型训练阶段,本实验使用的操作系统为 Ubuntu 18.04,训练 GPU 为 2060ti,显存容量 6144M,测试环境为 Python3.8,测试框架为 Pytorch1.9.0,使用 CUDA 版本 10.2.89 作为框架计算和 cuDNN8.2.1 的神经网络加速库。迭代批量大小为 16,总迭代次数设置为 200 次,初始学习率设置为 0.001。

## 3 结果

本文的所有结果利用单张图片的检测时间和视频检测的帧数作为模型推理速度的评价指标。采用准确率(accuracy)、召回率(recall)、平均精度(average precision, AP)和权重大小等指标衡量检测效果。准确率越高,分拣正确的鸽子蛋就越多;召回率越高,漏检的鸽子蛋就越少。相关指标计算公式如下:

$$Precision = \frac{T_p}{T_p + F_p} \quad (13)$$

$$Recall = \frac{T_p}{T_p + F_N} \quad (14)$$

$$AP = \int_0^1 P(R) dR \quad (15)$$

其中,  $T_p$  为真实的正样本数量,  $F_p$  为虚假的正样本数量,  $F_N$  为虚假的负样本数量。

### 3.1 融合 SE 注意力机制前后对比

本实验分别用未添加 SE 模块、添加默认 SE 模块、改进 SE 模块 3 个模型进行对比实验,将这 3 种模型分别命名为 YOLOV5, YOLOV5-SE 和 YOLOv5-改进 SE。检测试验结果见表 1。

表 1 YOLOv5 模型融合 SE 注意力机制前后的检测指标对比

模型(Models)	Precision	Recall	mAP@0.5
YOLOv5	0.832	0.911	0.949
YOLOv5-SE	0.917	0.929	0.921
YOLOv5-改进 SE	0.922	0.938	0.939

从表 1 可见, 融合了 SE 注意力的 YOLOv5 检测的准确率、召回率得到显著提升, 整体检测准确率(Precision)达到 91.7%, 提高了 8.5%; 召回率(Recall)达到了 92.9%, 提高了 1.8%; 在融合改进 SE 注意力机制后, 准确率和召回率的指标进一步提升, 分别达到了 92.2% 和 93.8%, 对于漏检问题有了较好的改善.

### 3.2 消融实验

为验证本实验对 YOLOv5 改进后的检测效果, 在原数据集上展开消融实验, 以判断每个改进点的有效性, 依次在原有模型上添加检测层和改进 SE 注意力机制, 进行逐项改进. 结果如表 2 所示(使用该项改进用“√”表示, 未使用该项改进用“—”表示).

表 2 消融实验各项指标对比

增加检测层	改进 SE 注意力模块	准确率	召回率	平均精度 mAP@0.5
—	—	0.832	0.911	0.949
√	—	0.906	0.942	0.960
—	√	0.922	0.938	0.939
√	√	0.980	0.973	0.983

可以看出, 引入了改进 SE 注意力机制后, 准确率提高了 9%, 召回率提升了 2.7%; 而增加了更小检测层之后, 准确率的提升虽然不如融合了改进 SE 注意力后那么惊艳, 但召回率却有了明显的提升, 高达 3.1%. 综合两种改进之后的新模型在准确率、召回率上都有了明显的提升, 其中准确率达到 98%, 提高了 14.8%; 召回率达到 97.3%, 提高了 6.2%. 实验结果表明, 对 YOLOv5 的每一项改进都取得了一定的效果, 综合改进网络结构和预测端的最终模型效果最佳. 部分检测效果见图 8.

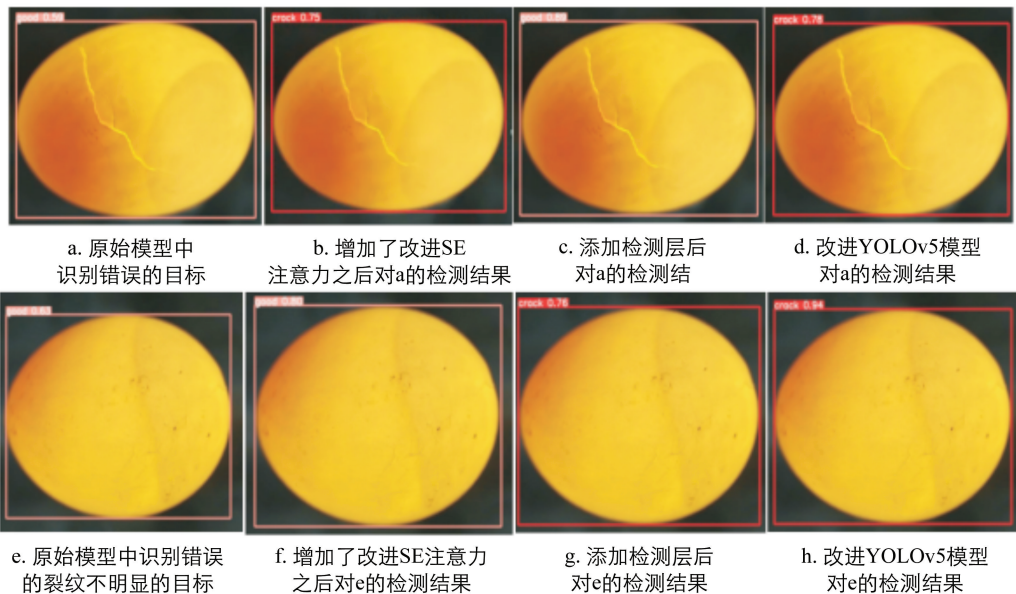


图 8 部分实际检测效果对比

### 3.3 不同模型检测效果比较

为了验证本研究实验所用算法的有效性和鲁棒性, 本文实验在相同的配置环境下, 用 Faster RCNN, YOLOv3-SPP, YOLOv5, 改进 YOLOv5 进行了基于同一数据集的对比实验, 结果见表 3.



表 3 Faster RCNN, YOLOv3-SPP, YOLOv5, 改进 YOLOv5 模型性能比较

模型 (Models)	Size (pixels)	Precision	Recall	FPS (帧/s)	Model size (MB)
Faster RCNN	640	0.492	0.926	6.6	193.2
YOLOv3-SPP	640	0.813	0.869	54	323
YOLOv5	640	0.832	0.911	125	3.4
改进 YOLOv5	640	0.980	0.973	111	4.6

结果显示, Faster-RCNN, YOLOv3-SPP, YOLOv5, 改进 YOLOv5 的准确率分别为 49.2%, 81.3%, 83.2%, 97.4%, 召回率分别为 92.6%, 86.9%, 91.1%, 97.1%, 本实验的改进 YOLOv5 在准确率和召回率方面均为最高水平. 虽然在模型的优化过程中, 增加了一定的计算量, 导致检测速度略有下降, 但差别不大, 且在 30 帧/s 的视频流下可以流畅运行并准确识别. 因此, 改进 YOLOv5 模型具有最优检测效果.

## 4 讨论

由表 1 可知, 融合了 SE 注意力机制的 YOLOv5 模型在各项指标上都明显优于原始的 YOLOv5 模型, 表明在 neck 的 C3 模块后面融合的 SE 注意力机制的确提高了对鸽蛋纹理特征的提取能力, 从而能够更好地保留裂纹特征, 提高了检测的准确率和召回率. 但平均池化层依旧会忽略掉一部分裂纹特征, 导致仍有相当一部分裂纹鸽蛋被错检和漏检. 而将 SE 注意力机制优化后, 通过反向的最大池化层提升了裂纹特征的权重; 同时, Hardswish 激活函数的全连接层替换 ReLU 激活函数的全连接层后, 在基本保证了原有推理速度的前提下, 避免了 ReLU 激活函数在输入为负时存在的神经元坏死问题, 确保神经网络误差正常反馈(图 9), 检测效果有了进一步提升.

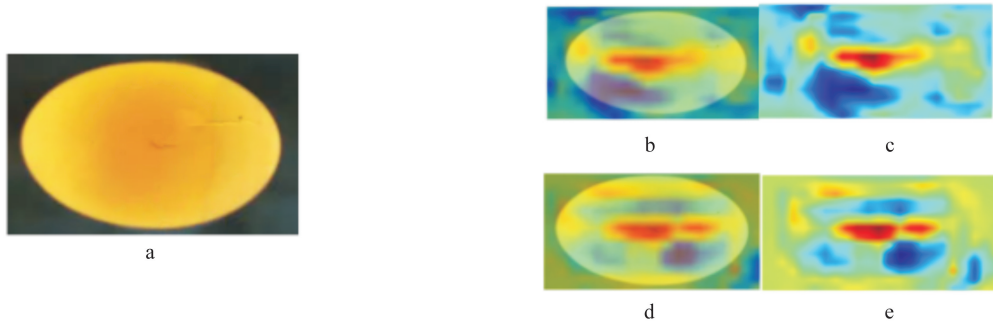


图 9 融合 SE 注意力机制和改进 SE 注意力机制特征图对比

由表 2 可知, 单项的改进虽然在整体上有所提升, 但却有着各自的缺陷: 增加了检测层的改进模型虽然提升了对细微裂纹的鸽蛋的检测效果, 但对于在 YOLOv5 模型下裂纹明显的错检鸽蛋的检测结果基本没有得到纠正; 而只融合了 SE 注意力机制的改进模型, 虽然在裂纹明显的鸽蛋的错检方面有了较为明显的改善, 但对 YOLOv5 模型下比较细微裂纹的错检或漏检并没有得到很好的纠正. 在融合两种改进之后, 模型在减少错检和漏检方面都表现出较好的效果. 增加检测层和融合注意力机制的代价是计算量增加, 导致检测速度降低了大约 11%, 但相对于细微裂纹检测精度接近 20% 和召回率接近 12% 的提升, 速度上的小幅降低是可以接受的. 因此, 这两种改进是比较有效的.

Faster RCNN 作为深度学习两阶段的目标检测算法, 算法原理是通过 RPN (Region Proposal Network) 网络先提取出物体的候选位置 (Region Proposal), 然后对候选位置进行物体分类、边界框回归和置信度计算, 因此, Faster RCNN 的召回率会高于 YOLO 系列算法. 从表 3 的实验结果中可以看出, Faster RCNN 的准确率是四种模型中最低的, 表明其较难准确分类出完好鸽蛋和裂纹鸽蛋. 同时, 两阶段算法的特性决定了 Faster RCNN 的运行速度要低于 YOLO 系列算法. YOLOv3-SPP 和 YOLOv5 的差别不大, 但 YOLOv5 在 neck 部分引入 CSP 结构, 加强了网络特征的融合能力, 从而提高了模型的效果. 此外, YOLOv5 模型在检测速度和模型大小上优于 YOLOv3-SPP 模型, 模型性能在整体上也略优于 YOLOv3-SPP 模型. 而改进的 YOLOv5 模型不仅性能上优于 YOLOv5, 而且继承了其占用内存资源较少、适合部署

于可移动嵌入式设备平台的优点, 相比其他 3 种模型都有较为明显的优势. 因此, 在鸽子蛋的品质分拣中选择 YOLOv5 算法, 并且在两方面进行改进优化, 目的是提高模型检测精度和模型特征提取能力:

1) 针对部分破损鸽蛋裂纹比较细微的特点, 本研究采用增加检测层的方式增强对细微裂纹特征的提取, 以尽可能减少错检和漏检.

2) 针对破损鸽蛋的裂纹特征提取难度比较大的特点, 在 neck 的特征提取模块后加入改进的 SE 注意力机制, 以增强裂纹特征的权重和模型的特征提取能力, 进一步提高模型检测精度.

本研究自主设计的检测设备(图 7)对鸽子蛋进行分拣, 对于裂纹位于蛋身上的鸽蛋有了较好的检出效果, 但对鸽蛋锐端和钝端的细微裂纹, 因为角度的原因, 无法完全采集并识别到, 因此, 我们假设鸽蛋锐端和钝端处没有裂纹.

## 5 结论

本研究实验针对鸽蛋破损的检测任务, 提出了一种基于改进 YOLOv5 的裂纹检测方法, 通过数据增强扩大数据集、增加检测层、在主干网络中添加注意力机制并对注意力机制的全局平均池化层和全连接层进行改进, 有效提高了裂纹鸽蛋的检测准确率.

实验结果表明, 本次实验构建的基于改进 YOLOv5 的裂纹鸽蛋的分拣识别模型整体性能优于 Faster RCNN、YOLOv3-SPP, 对于裂纹鸽蛋的识别准确率高、速度快. 在检测设备上进行检测时, 模型对鸽子蛋壳裂纹的检测精度达到了 98%, 召回率达到了 97.3%, 对单张图片的检测速度在 0.09 s 以内, 开发板环境下可在帧数约 30 FPS 的实时检测中流畅地运行. 该实验模型已经可以满足企业在鸽子蛋品质分拣上的应用. 并且模型大小仅为 4.6 Mb, 可搭载在开发板、移动设备等轻量级的设备上, 是一个能够大规模部署的轻量级模型, 可以为搭建鸽子蛋品质检测流水线提供良好的技术支撑.

## 参考文献:

- [1] ARIVAZHAGAN S, SHEBIAH R N, SUDHARSAN H, et al. External and Internal Defect Detection of Egg Using Machine Vision [J]. *Journal of Emerging Trends in Computing and Informator Science*, 2013, 4(3): 257-261.
- [2] 邓海霞, 刘友明, 文友先, 等. 基于机器视觉群体鸡蛋尺寸的检测方法 [J]. *华中农业大学学报*, 2006, 25(4): 452-454.
- [3] 潘磊庆, 屠康, 刘鹏, 等. 影响鸡蛋声学响应因素的分析研究[C]// 第八届中国蛋品科技大会论文集. 2009: 516-522.
- [4] 马卫东, 于亚萍, 张樱飘, 等. 一种声音信号判断鸡蛋裂缝系统设计 [J]. *机电工程技术*, 2020, 49(11): 53-54, 118.
- [5] 贺静, 王树才. 基于 DSP 实时图像分割算法的鸡蛋蛋壳破损检测 [J]. *湖南科技学院学报*, 2010, 31(4): 55-58, 82.
- [6] 王树才, 任奕林, 陈红, 等. 利用敲击声音信号进行禽蛋破损检测和模糊识别 [J]. *农业工程学报*, 2004, 20(4): 130-133.
- [7] 邓颖, 吴华瑞, 孙想. 基于机器视觉和穿戴式设备感知的村镇老年人跌倒监测方法 [J]. *西南大学学报(自然科学版)*, 2021, 43(11): 186-194.
- [8] 杜文飞, 李春光, 万四海. 管道漏磁检测的智能方法综述 [J]. *西南师范大学学报(自然科学版)*, 2022, 47(6): 1-7.
- [9] 涂伟沪, 蔡玲霞, 李学军. 基于改进蝗虫算法优化 Canny 算子的鸡蛋裂纹图像检测 [J]. *食品与机械*, 2022, 38(2): 167-172, 202.
- [10] 赵祺喜, 罗阳帆, 黄杏彪, 等. 基于机器视觉和 YOLOv4 的破损鸡蛋在线检测研究 [J]. *现代农业装备*, 2022, 43(1): 8-16.
- [11] 汤青萍, 常玲玲, 卜柱, 等. 鸽蛋壳的物理性状检测及超微结构观察与分析 [J]. *青岛农业大学学报(自然科学版)*, 2019, 36(1): 41-44, 74.
- [12] 万卓, 叶明, 刘凯. 基于改进 YOLOv4 的电机端盖缺陷检测 [J]. *计算机系统应用*, 2021, 30(3): 79-87.
- [13] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [14] 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. IEEE, 2016: 779-788.
- [15] 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.
- [16] DONG S, MA Y H, LI C M. Implementation of Detection System of Grassland Degradation Indicator Grass Species Based on YOLOv3-SPP Algorithm [J]. Journal of Physics: Conference Series, 2021, 1738(1): 012051.
- [17] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a New Backbone that can Enhance Learning Capability of CNN [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). June 14-19, 2020, Seattle, WA, USA. IEEE, 2020: 1571-1580.
- [18] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [19] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 8759-8768.
- [20] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述 [J]. 计算机学报, 2017, 40(6): 1229-1251.
- [21] ZAKI P W, HASHEM A M, FAHIM E A, et al. A Novel Sigmoid Function Approximation Suitable for Neural Networks on FPGA [C]//2019 15th International Computer Engineering Conference (ICENCO). December 29-30, 2019, Cairo, Egypt. IEEE, 2020: 95-99.
- [22] MERCIONI M A, HOLBAN S. Soft-Clipping Swish: a Novel Activation Function for Deep Learning [C]//2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI). May 19-21, 2021, Timisoara, Romania. IEEE, 2021: 225-230.
- [23] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for Activation Functions [EB/OL]. 2017: arXiv: 1710. 05941. <https://arxiv.org/abs/1710.05941>.
- [24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27-November 2, 2019, Seoul, Korea (South). IEEE, 2020: 1314-1324.
- [25] GAO S, WU J M, DA C, et al. SATB-Nets: Training Deep Neural Networks with Segmented Asymmetric Ternary and Binary Weights [C]//International Conference on Neural Information Processing. Cham: Springer, 2018: 700-710.
- [26] HENDRYCKS D, GIMPEL K. Bridging Nonlinearities and Stochastic Regularizers with Gaussian Error Linear Units [J]. CoRR, 2016, abs/1606.08415: 1-10.

责任编辑 潘春燕