doi:10.6041/j.issn.1000-1298.2025.05.046

基于 OD_SeGAN 的断奶前仔猪实例分割方法

李　鹏[1,2]　沈明霞[1,2]　刘龙申[1,2]　陈金鑫[2]　薛鸿翔[2]　衡　熙[2]　孙玉文[2]
(1.南京农业大学人工智能学院,南京 210031;2.农业农村部养殖装备重点实验室,南京 210031)

摘要:在猪只智慧养殖中,猪只实例分割方法是实现猪只自动化检测的关键技术之一,但在实际分割场景中,存在猪只遮挡粘连等现象,易导致分割困难。针对产房中仔猪分割困难问题,本文提出一种基于 YOLO v5s 和 GAN(Generative adversarial network)的实例分割模型 OD_SeGAN。该方法通过目标检测算法 YOLO v5s 提取出仔猪目标,并输入至语义分割算法 GAN 实现分割,并使用空洞卷积替换 GAN 中的普通卷积,扩大网络感受野;其次,使用挤压-激励注意力机制模块,增强模型对仔猪全局特征的学习能力,提高模型的分割精度。实验结果表明,OD_SeGAN 在测试集上 IoU 为 88.6%,分别比 YOLO v5s_Seg、Cascade_Mask_RCNN、Mask_RCNN、SOLO、Yolact 高 3.4、3.3、4.1、9.7、8.1 个百分点。将 OD_SeGAN 应用于仔猪窝均质量估测任务中,测得仔猪窝均质量和仔猪像素点数之间皮尔逊相关系数为 0.956。OD_SeGAN 在实际生产场景中具有良好的仔猪分割性能,可为仔猪窝均质量估测等后续研究提供技术基础。

关键词:仔猪;实例分割;注意力机制;GAN;窝均质量
中图分类号:TP18;S24　　**文献标识码**:A　　**文章编号**:1000-1298(2025)05-0482-10　　**OSID**:

Instance Segmentation Method of Pre-weaning Piglets Based on OD_SeGAN

LI Peng[1,2]　SHEN Mingxia[1,2]　LIU Longshen[1,2]　CHEN Jinxin[2]　XUE Hongxiang[2]　HENG Xi[2]　SUN Yuwen[2]
(1. *College of Artificial Intelligence*, *Nanjing Agricultural University*, *Nanjing* 210031, *China*
2. *Key Laboratory of Breeding Equipment*, *Ministry of Agriculture and Rural Affairs*, *Nanjing* 210031, *China*)

Abstract: In the research on smart pig breeding, the pig instance segmentation method is one of the key technologies to realize automatic detection of pigs. However, in actual segmentation scenarios, there is occlusion and adhesion phenomenon, which makes pig segmentation difficult. Aiming at the difficulty of piglet segmentation in the farrowing room, an instance segmentation model OD_SeGAN was proposed based on YOLO v5s and generative adversarial network(GAN). This method extracted the piglet target through the target detection algorithm YOLO v5s, and inputed it into the semantic segmentation algorithm GAN to achieve segmentation, and used dilated convolution to replace the ordinary convolution in GAN to expand the network receptive field; secondly, a squeeze-incentive attention mechanism was used module to enhance the model's ability to learn the global characteristics of piglets and improve the model's segmentation accuracy. Experimental results showed that OD_SeGAN's IoU on the test set was 88.6%, which was 3.4, 3.3, 4.1, 9.7, and 8.1 percentage points higher than YOLO v5s_Seg, Cascade_Mask_RCNN, Mask_RCNN, SOLO, and Yolact, respectively. OD_SeGAN was applied to the piglet litter average weight estimation task, and the Pearson correlation coefficient between the piglet litter average weight and the number of piglet pixels was measured to be 0.956. The OD_SeGAN proposed had good piglet segmentation performance in actual production scenarios, and can provide a technical basis for subsequent research such as piglet litter weight estimation.

Key words: piglets; instance segmentation; attention mechanism; GAN; average litter weight

收稿日期:2024-03-11　修回日期:2024-04-17
基金项目:国家重点研发计划项目(2021YFD2000800)
作者简介:李鹏(1997—),男,博士生,主要从事精准养殖和机器视觉研究,E-mail:lipeng@stu.njau.edu.cn
通信作者:沈明霞(1964—),女,教授,博士生导师,主要从事机器视觉和信息农业化研究,E-mail:mingxia@njau.edu.cn

0　引言

计算机视觉技术凭借其无接触式、鲁棒性强等特点，成为实现猪场自动化养殖、评价猪只福利的重要技术手段之一[1-2]。近年来，随着电子技术的发展，算力成本大幅降低，研究者将计算机视觉技术引入猪场实际生产应用中，在猪只个体识别[3]、发情检测[4]、母猪分娩[5]、计数[6]、姿态行为分类[7-9]及体尺体质量测量[10-12]等任务中，取得了良好效果。

在仔猪计数、体质量体尺测量、姿态行为分类等任务中，实例分割方法能够提供比目标检测方法更加精准的位置和语义信息，在多个研究领域中得到广泛应用[13]，文献[14-19]表明实例分割技术可作为仔猪智能化管理的基础技术。断奶前仔猪具有生长速度快、适应能力差、免疫力低等特点，断奶前仔猪窝均质量不仅可以作为衡量母猪母性的重要指标，且仔猪断奶质量与腹泻率和死亡率呈显著负相关[20]，因此养殖人员需要及时关注断奶前仔猪体质量及变化，利用实例分割方法估测仔猪窝均质量，可以保证实时检测并获得较高准确率，极大地减少人员工作量。

为解决在复杂生产环境下，断奶前仔猪存在遮挡粘连等导致仔猪分割困难的问题，本文构建基于目标检测网络 YOLO v5s[21] 和生成对抗网络（Generative adversarial network，GAN）的仔猪分割模型 OD_SeGAN，首先通过改进损失函数的 YOLO v5s 提取仔猪子图像以降低后续分割的空间复杂度，将子图像输入使用空洞卷积并增加 SE 模块的 GAN 实现仔猪分割，并将 OD_SeGAN 模型应用于无接触式断奶前仔猪窝均质量估测研究中，以期为仔猪窝均质量测量任务提供自动化、无接触式实现方法。

1　实验材料

1.1　数据采集

实验数据采集于江苏省丹阳市司徒镇银光猪场。数据采集系统主要由摄像头、网络传输模块与数据存储模块 3 部分组成，如图 1 所示，摄像头采集仔猪视频数据，经交换机传输至录像机进行本地存储，数据经交换机和一级路由器上传至服务器，可通过服务器端实时查看。选择 14 栏共 144 只仔猪作为实验目标，每栏有 8～12 头仔猪和 1 头母猪，分娩栏长×宽为 2.45 m×1.8 m，每栏正上方 2.2 m 处配有 1 个分辨率为 1 920 像素×1 080 像素的海康威视 DS-2CD3325-I 型摄像头，摄像头速率为 25 f/s，对分娩栏中的仔猪进行俯拍。采集的仔猪日龄在 9～21 d 之间，且视频均采集于 08:00—16:00，采集时间为 2022 年 8 月 1—12 日。

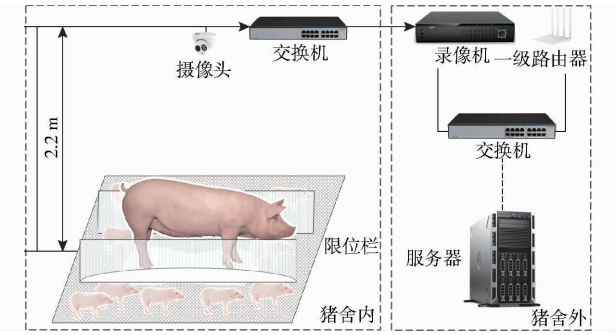

图 1　系统示意图

Fig. 1　Sketch of video acquisition

仔猪窝均质量数据采自 2022 年 8 月 11 日 10:00—14:00 间，共获取 13 栏仔猪窝均质量数据，窝质量范围为 31.3～68.38 kg，窝均质量范围为 3.91～6.22 kg。

1.2　数据集制作

在采集的视频中每天随机抽取 1 h 视频进行截取，每 20 s 截取一幅图像，剔除相似度过高的图像，共获得 4 160 幅有效图像。在产房中，仔猪存在严重遮挡（仔猪被母猪、限位栏以及其他仔猪遮挡）。在本研究中，遮挡被分为 4 种：身体分离遮挡，即仔猪身体被分成两个或多个部分；身体部分缺失遮挡，即仔猪的末端部分（例如头部、尾部）被遮挡；身体分离遮挡和身体部分遮挡，即同时存在身体分离遮挡和身体部分遮挡；全部遮挡，即仔猪身体被全部遮挡。每只仔猪只存在一种遮挡情况，仔猪具体遮挡情况见表 1。

本文标注的实例分割数据集标注内容包含非完全遮挡的仔猪，使用开源软件 labelme 标注，且标签只有仔猪一类，对于不可预测的仔猪遮挡部分（即仔猪身体缺失严重部分）不予标注。在下述实验中，将 1～4 栏的仔猪作为测试集，5～14 栏的仔猪作为训练集，测试集和训练集比例为 0.28∶0.72。

2　方法

2.1　实验流程

本文提出的仔猪分割模型由一个目标检测模型 YOLO v5s 和一个语义分割模型 GAN 构成，将实例分割数据集分成目标检测数据集和语义分割数据集，利用标注好的仔猪轮廓信息生成能包含该轮廓的最小外接矩形框生成目标检测数据集，将标注好的仔猪轮廓信息和对应的仔猪单独分割出来生成语义分割数据集，且不改变原本数据集的划分方式；利用生成的两组数据集训练 OD_SeGAN，并与 YOLO v5s_Seg 等实例分割模型进行模型性能对比，选择精度高的模型应用于仔猪窝均质量估测任务中，并进行结果分析。实验流程如图 2 所示。

表 1 仔猪遮挡情况说明

Tab. 1 Description of piglet occlusion

具体遮挡情况	产床编号			总计
	1 ~ 4 栏	5 ~ 9 栏	10 ~ 14 栏	
可见仔猪数量	11 004(94%)	14 365(96%)	15 757(94%)	41 126(95%)
身体分离遮挡下仔猪数量	1 140(10%)	1 414(9%)	1 573(9%)	4 127(9%)
身体部分遮挡下仔猪数量	2 053(18%)	2 751(18%)	3 603(22%)	8 407(19%)
身体分离遮挡和身体部分遮挡下仔猪数量	2 428(21%)	3 784(25%)	3 510(21%)	9 722(22%)
无遮挡下仔猪数量	5 383(46%)	6 416(43%)	7 071(42%)	18 870(44%)
完全遮挡数量	709(6%)	570(4%)	991(6%)	2 270(5%)
图像数量	1 171	1 409	1 580	4 160
仔猪数量	40	52	52	144

注:可见仔猪是指俯视情况下可见仔猪,包括被遮挡仔猪及无遮挡仔猪。

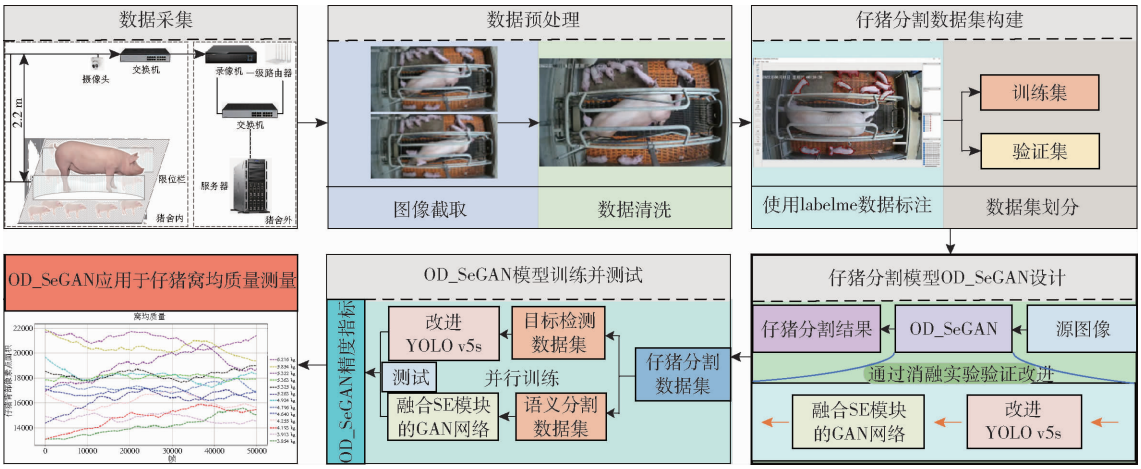

图 2 实验流程图

Fig. 2 Overall flowchart of test

2.2 OD_SeGAN 模型构建

2.2.1 主体框架

OD_SeGAN 模型实现仔猪实例分割主要分为两个过程,首先仔猪目标检测部分使用 YOLO v5s 目标检测网络提取出完整的仔猪,在 YOLO v5s 识别的基础上增加识别框的宽度、高度,值为 10 个像素,对于处于源图像边缘的识别框不做处理,然后将提取出的仔猪子图像缩放至 256 像素×256 像素并输入至下一部分网络。提取子图像的做法创建了用于预测仔猪掩膜的子空间,该空间的复杂度远低于源图像,有利于提高模型整体的精确度;其次,仔猪分割部分使用 GAN 实现仔猪分割,GAN 由一个生成器和判别器组成,生成器结构基于 Unet 结构设计,用于生成仔猪分割图像,判别器由 5 层 CNN 组成,用于判别生成器所生成分割图像的准确度。OD_SeGAN 模型分割流程如图 3 所示。

GAN 使用普通卷积作为特征提取方法,由于卷积是局部运算,获取全局信息的能力不足,扩大感受野会在一定程度上提高模型的精度。普通卷积扩大感受野只能通过增大卷积核大小或者增加模型层

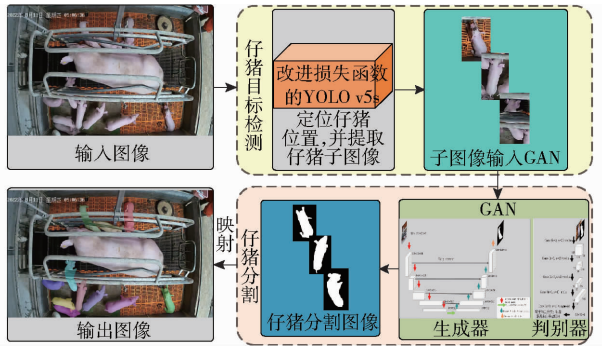

图 3 OD_SeGAN 模型算法流程图

Fig. 3 OD_SeGAN model algorithm process

数,但会增加模型参数量。本文在生成器部分,通过引入空洞卷积替换常规卷积来增加感受野,并在中间层添加挤压-激励注意力模块来增加模型获取仔猪全局特征的能力。

2.2.2 空洞卷积

在仔猪分割任务中,往往存在不同程度的遮挡情况,遮挡严重阻碍分割模型对仔猪的个体识别准确率,本文通过增加网络感受野的方式增加模型的识别准确率,通过获取较大范围的特征信息,网络可以更好地推断被遮挡仔猪的完整形态。文献[22]

提出的空洞卷积结构可以在不增加模型参数量以及深度的情况下增加单次卷积感受野，使每个卷积都能输出包含较大区域的信息，同时可以保证输出的特征向量长度保持不变。本文将生成器下采样过程中的普通卷积均换成空洞卷积，空洞卷积示意图如图 4 所示。

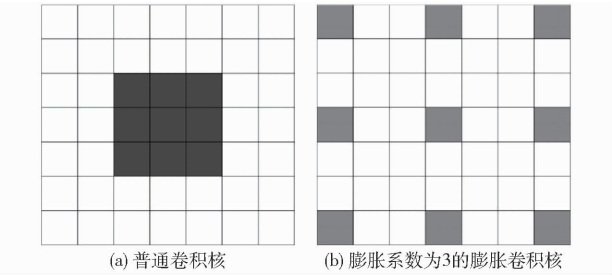

图 4　空洞卷积示意图

Fig. 4　Schematics of atrous convolution

空洞卷积感受野计算公式为

$$F = 2(d-1)(k-1)+k \tag{1}$$

式中　k——输入的卷积核尺寸

d——膨胀系数

F——空洞卷积感受野大小

2.2.3　挤压-激励注意力模块

在仔猪分割任务中，空洞卷积在不增加模型参数的情况下扩大了局部感受野，但是不能形成目标仔猪的全局特征信息。为此本文引入挤压-激励自注意力机制[23]（Squeeze - excitation，SE），以获得仔猪的全局特征关系，减小遮挡对分割模型效果的影响。SE 模块网络结构在 GAN 中的应用如图 5 所示。

SE 模块通过挤压和激励两个核心部分实现获取仔猪全局特征信息的过程。在挤压阶段，该模块将来自各个通道的空间特征信息聚合并转化为具有全局感受野的全局特征，将每个通道上特征图的空间维度降低为一个标量（单个数值），从而获取每个通道上的全局信息。对于给定的特征图，其维度为 $H \times W \times C$，其中 H、W、C 分别表示特征图的高度、宽度、通道数。挤压操作对每个通道的特征图进行全

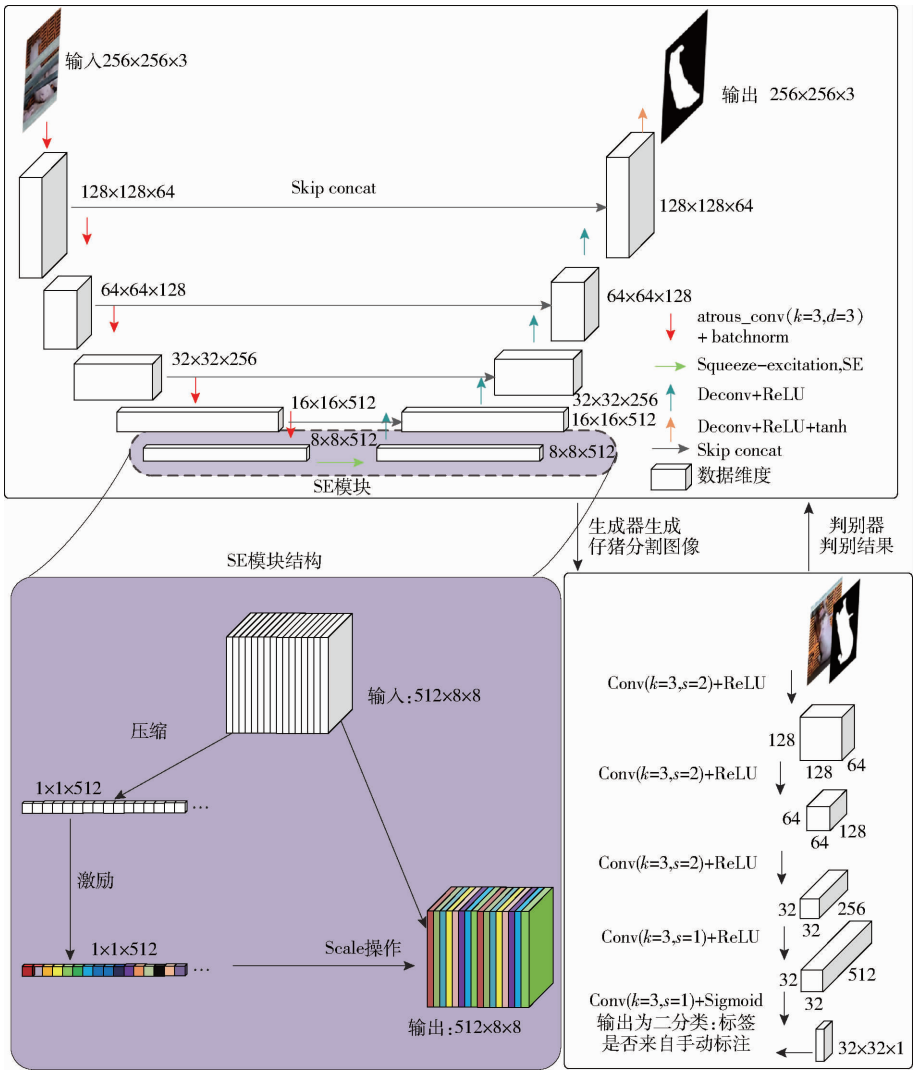

图 5　SE 模块在 GAN 中的结构图

Fig. 5　Structure diagram of SE module in GAN

局平均池化,生成一个通道描述符 z,通道描述符上第 c 个元素 z_c 计算式为

$$z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \tag{2}$$

式中 u_c——特征图中第 c 个通道特征值

(i,j)——特征图空间位置的坐标

在激励阶段,激励操作利用挤压得到的通道描述符 z 来学习每个通道的权重。该过程通过一个门控机制实现,门控机制包括 2 个全连接层和 1 个激活函数。第 1 个全连接层降低维度,第 2 个则恢复到原始通道数,可以增加模型的非线性并减少参数量。权重向量 $\boldsymbol{s}$ 计算式为

$$\boldsymbol{s} = \sigma(W_2\delta(W_1 z)) \tag{3}$$

式中 σ——Sigmoid 激活函数,确保权重在 0~1,允许模型进行特征重标定

δ——ReLU 激活函数

W_1、W_2——全连接层权重

激励得到的权重向量 $\boldsymbol{s}$ 用于对原始特征图进行缩放操作(按元素相乘),以得到重新校准的特征图,新特征图第 c 个通道计算式为

$$\widetilde{u}_c = s_c u_c \tag{4}$$

式中 $\widetilde{u}_c$——经过重新校准的第 c 个通道特征图

s_c——对应的通道权重

SE 模块具有即插即用的优点,本文在生成器部分融合 SE 模块,即在生成器中间部分插入 SE 模块,SE 模块接受的特征图维度为 8 ×8 ×512,经过挤压、激励和校准后生成包含仔猪全局特征的新特征图,维度不变。

2.2.4 YOLO v5s 损失函数改进

YOLO v5s 目标检测模型用于第 1 阶段提取仔猪图像,其准确度直接影响整体分割的准确度。损失函数是衡量模型预测准确度的一个重要指标,同时影响训练过程中模型的收敛速度。目标检测模型损失函数包含 3 部分,即置信度损失、分类损失以及定位损失。置信度损失采用二元交叉熵损失函数(BCE Loss),用于衡量模型对仔猪个体存在的置信程度以及是否成功检测到物体;分类损失采用 BCE Loss,用于衡量模型对检测到的目标进行分类的准确性;定位损失采用高效交并比损失函数(EIoU Loss),用于度量模型对仔猪目标位置的预测精度,包括中心坐标和宽高信息。本文研究对象仅为仔猪目标,不对仔猪以外的其他目标做识别,故分类损失为 0。YOLO v5s 整体损失函数为

$$L_{oss} = \lambda_1 L_{obj} + \lambda_2 L_{Bbox} \tag{5}$$

式中 L_{obj}——置信度损失

λ_1、λ_2——对应系数,取 1/2

L_{Bbox}——定位损失

完整交并比损失(CIoU Loss)作为 YOLO v5s 原损失函数,使用纵横比描述相对值,且宽和高的损失高度耦合,CIoU 损失函数为

$$L_{CIOU} = L_{IOU} + L_{dis} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{(w^c)^2 + (h^c)^2} \tag{6}$$

式中 L_{IOU}、L_{dis}——交并比损失和距离损失

$\rho^2(b,b^{gt})$——真实框和预测框中心点的欧氏距离

w^c、h^c——预测框和真实框的最小外接框宽度和高度

IoU——预测框和真实框的并集面积和交集面积的比

本文使用 EIoU 损失函数代替 YOLO v5s 中使用的 CIoU 损失函数,进一步考虑了仔猪目标的尺寸信息,将预测框和真实框纵横比的影响因子进行解耦,分别计算预测框和真实框的长和宽,可以更精确地描述不同边界框之间的重叠情况,从而提高目标定位的准确性。EIoU 损失为

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w,w^{gt})}{(w^c)^2} + \frac{\rho^2(h,h^{gt})}{(h^c)^2} \tag{7}$$

式中 L_{asp}——高度和宽度损失

w、h——检测框的宽度和高度

2.2.5 GAN 损失函数设计

GAN 在训练过程中是通过生成器和判别器相互对抗提高仔猪分割效果,生成器和判别器网络是独立存在的,其中判别器的任务是对掩膜图像进行分类,判断其是否为生成器生成的掩膜图像,本文使用真实数据和生成数据的交叉熵损失作为判别器的损失函数,生成器生成输入图像的仔猪掩膜图像,属于语义分割任务,且需要联合判别器进行训练,本文使用 Dice Loss 作为生成器的损失函数,加上 L1 正则项,以防止模型的过拟合问题。GAN 损失函数为

$$\begin{cases} generator_loss = \lambda_1\left[1 - \dfrac{1}{n}\dfrac{2y \cap y'}{y \cap y + y' \cap y'} + \lambda_2\dfrac{1}{n}(|y - y'|)\right] + \lambda_3\dfrac{1}{n}(-\lg(D(y'))) \\ discriminator_loss = -\dfrac{1}{n}((\lg(D(y)) + \lg(1 - D(y')))) \end{cases} \tag{8}$$

式中 y——真实标签

y'——生成器对仔猪图像的映射

n——样本数

$D(y)$——判别器对真实标签的判别结果

β_1、β_2、β_3——各部分损失在生成器损失函数中比重，本实验中设置 $\beta_1 = \beta_2 = \beta_3 = 1$

2.3　仔猪窝均质量度量方式

仔猪作为一种非刚性个体，其在图像中的表现形式与姿态多样，且变化幅度大，通过提取其体尺特征（体宽、体长等）回归窝均质量较为困难，因此本文使用俯拍视角下仔猪分割轮廓内像素点数（以下称仔猪像素点数）衡量仔猪窝均质量。将 OD_SeGAN 应用于仔猪窝均质量的估测任务中，使用 OD_SeGAN 的分割结果计算仔猪像素点数并分析仔猪窝均质量和平均每头仔猪像素点数之间的皮尔逊相关性。

皮尔逊相关系数（Pearson，r）可以衡量仔猪窝均质量和仔猪像素点数之间线性相关程度，取值范围为[-1,1]，r 绝对值越大，表示两个变量间的线性相关程度越强。

3　结果与分析

3.1　模型训练与测试

3.1.1　实验平台搭建

模型训练依赖于 Pytorch 1.8.1 深度学习框架，实验平台处理器型号为 Intel ® CoreTM i9 - 10900k，显卡型号为 NVIDIA GTX3090，显卡内存容量为 24 GB，在 Ubuntu 20.04 操作系统上配置 Python 3.9、CUDA 11.1、Opencv 4.5.1 等深度学习环境。

3.1.2　评价指标

采用精确率（Precision）、召回率（Recall）、IoU 和单帧视频平均检测时间作为实例分割模型评价性能指标。在消融实验时，以 mAP@0.5 作为目标检测模型 YOLO v5s 的评价精度指标，以 IoU 作为语义分割模型 GAN 的评价精度指标。

3.1.3　消融实验

在训练目标检测模型 YOLO v5s 时，以 EIoU Loss 代替 CIoU Loss，以期获得更好的模型精度。在本实验中，训练集来自第 5 ~ 14 栏所采集的图像，采用 Mosaic 数据增强技术[24]，并将图像输入尺寸设置为 640 像素×640 像素，优化器采用随机梯度下降法（Stochastic gradient descent，SGD），初始学习率设为 0.01，动量初始值 0.95，权重衰减系数为 0.001，批次大小为 16，训练轮次设为 160。

为了验证优化模型的损失函数是否可以改进模型精度，在消融实验中，对比 CIoU、DIoU[25]、EIoU[26]、GIoU[27] 及 SIoU[28] 损失函数，其损失值变化如图 6 所示，其对应的模型精度如表 2 所示。

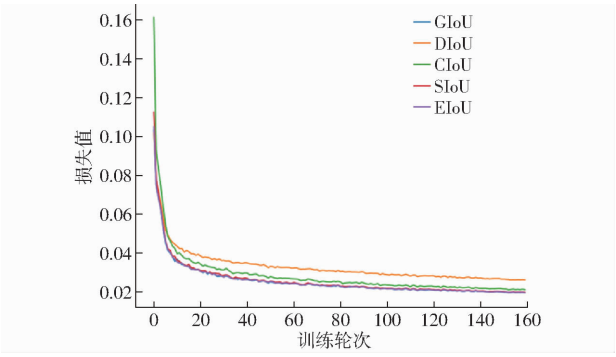

图 6　不同损失函数对应的损失值曲线

Fig. 6　Loss curves corresponding to different loss functions

表 2　不同损失函数对应的 mAP@0.5 和损失值

Tab. 2　The mAP@0.5 and loss values corresponding to different loss functions

模型	mAP@0.5/%	损失值
YOLO v5s - GIoU	98.2	0.020 105
YOLO v5s - DIoU	98.1	0.035 345
YOLO v5s - CIoU	98.4	0.024 204
YOLO v5s - SIoU	98.4	0.022 215
YOLO v5s - EIoU	98.6	0.019 481

由图 6 可知，模型的损失在 120 轮次时基本稳定，选取第 120 轮的损失值和 mAP@0.5 进行分析。对比 EIoU 损失函数和原损失函数 CIoU 的损失曲线，使用 EIoU 损失函数的模型收敛速度更快；与其他损失函数模型相比，使用 EIoU 损失函数的模型达到最高的 mAP@0.5 值（98.6%）和最低的损失值（0.019 481），故选用 EIoU 损失函数作为 YOLO v5s 的损失函数。

为验证本文提出的 GAN 模型的高精度，在训练 GAN 网络模型时，通过将普通卷积替换成空洞卷积以及添加挤压-激励自注意力机制模块获得不同的效果。在训练过程中，优化器采用自适应矩估计梯度下降算法（Adaptive moment estimation，Adam），学习率设置为 0.000 2，批次大小为 16，共训练 160 轮，消融实验结果如表 3 所示。

表 3　消融实验结果

Tab. 3　Results of ablation experiment　%

空洞卷积模块	挤压-激励注意力模块	IoU
×	×	89.8
✓	×	90.0
×	✓	92.4
✓	✓	92.6

注：✓表示 GAN 中采用该模块，×表示未采用该模块。

由表 3 可知，本文提出的 GAN 模型达到最高精度，IoU 为 92.6%。图 7 展示了仔猪图像在训练过程中随着训练轮次变化的部分可视化结果。由图 7 可知，当断奶前仔猪存在遮挡时，GAN 可以生成较

为准确的掩膜图像。

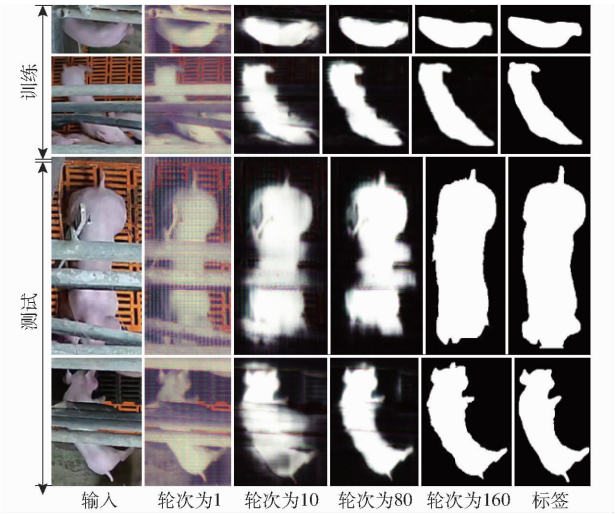

图 7 可视化结果展示

Fig. 7 Visual results display

消融实验结果表明,通过优化 YOLO v5s 损失函数提高了 YOLO v5s 仔猪目标检测精度,改进 GAN 模型结构提高了 GAN 的仔猪分割精度,不同部分改进对 OD_SeGAN 检测结果影响如表 4 所示。

表 4 不同部分改进对 OD_SeGAN 检测结果影响

Tab. 4 Impact of different parts of improvement on OD_SeGAN detection results

YOLO v5s 损失函数	GAN 结构	精确率/%	召回率/%	IoU/%
×	×	97.7	96.8	86.8
✓	×	97.2	97.6	87.4
×	✓	97.7	96.9	87.8
✓	✓	97.2	97.7	88.6

注:✓表示 OD_SeGAN 中采用该模块,×表示未采用该模块。

3.1.4 分割模型效果对比

为了进一步验证本文所提 OD_SeGAN 模型的高精度性能,在相同数据集上面,与 YOLO v5s_Seg、Cascade _ Mask _ RCNN[29]、Mask _ RCNN[16]、Yolact[30]、SOLO[31]等实例分割模型进行效果对比,结果对比如表 5 所示,n 为源图像中仔猪数量,OD_SeGAN 检测速度受到源图像中仔猪数量影响。

表 5 不同模型对比结果

Tab. 5 Comparison results of different models

模型	骨干网络	精确率/%	召回率/%	IoU/%	单帧视频平均检测时间/ms
YOLO v5s_Seg		97.2	95.1	85.2	11
Cascade_Mask_RCNN	Convnext	96.9	96.3	85.3	81
Mask_RCNN	Resnet50	97.0	94.6	84.5	74
SOLO	Resnet50	79.6	80.5	78.9	35
Yolact	Resnet50	92.5	87.5	80.5	13
OD_SeGAN(本文)		97.2	97.7	88.6	$8+2.1n$

由表 5 可知,OD_SeGAN 模型在精度上高于其他 5 种算法。由于 OD_SeGAN 模型首先通过 YOLO v5s 提取出源图像中所有仔猪,然后将仔猪一一输入至 GAN 生成器中进行分割,导致模型在检测速度上仅优于 Cascade _ Mask _ RCNN、Mask _ RCNN,和 SOLO 持平。

图 8 为表 5 中 6 种实例分割模型分割结果,其中区域 b1、b4 为身体部分遮挡,区域 b2 为身体分离遮挡和身体部分遮挡,区域 b3 为身体分离遮挡。针对产房中存在的几种遮挡问题,本文选取 3 幅图像作测试,并对图中 b1、b2、b3、b4 处进行分析。由图 8 可知,SOLO 在 b2 处存在重复识别问题,在 b4 处存在识别不全问题,Cascade_Mask_RCNN、Mask_RCNN、Yolact 和 YOLO v5_Seg 在 b2 处均未能体现两只仔猪的重叠效果。OD_SeGAN 在 b1、b2、b3、b4 处均能生成贴合仔猪边缘的掩膜,在 b2 处体现出仔猪的重叠现象。对于图中不存在遮挡情况的仔猪,6 种实例分割模型均能实现准确分割。

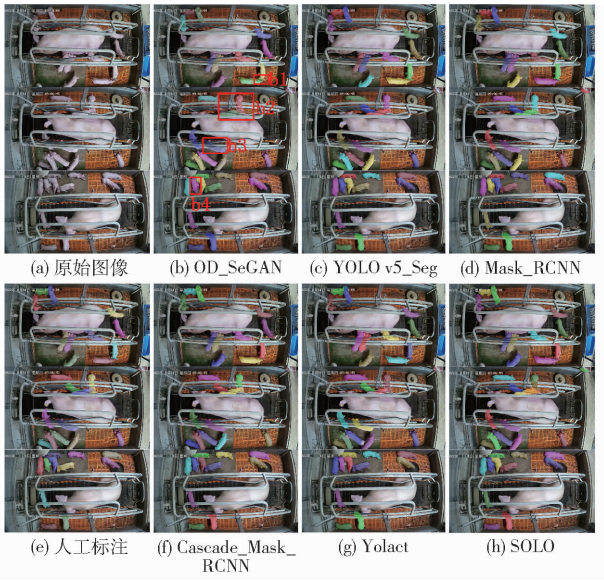

图 8 不同模型的实例分割结果

Fig. 8 Example segmentation results of different models

3.2 基于 OD_SeGAN 的无接触式仔猪窝均质量估测

3.2.1 仔猪像素点数获取与分析

在同一个栏位中,由于仔猪活动、母猪姿态转换等因素,导致不同时间点的仔猪像素点数存在变化。使用 OD_SeGAN 检测结果计算每栏平均每头仔猪像素点数,并衡量仔猪窝均质量。对每栏视频进行取帧,保留不存在完全遮挡的图像,每栏选取 50 000 幅图像按照时间顺序作为横坐标,平均每头仔猪像素点数与仔猪窝均质量的对应关系如图 9 所示。

由图 9 可知,随着仔猪在产房中活动,仔猪像素

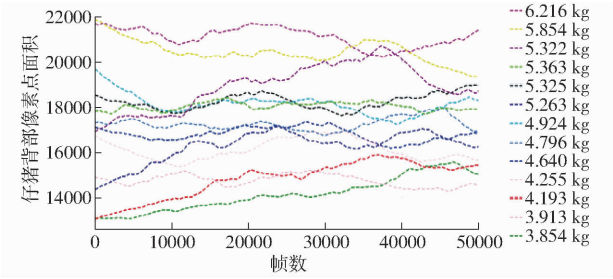

图 9　平均每头仔猪像素点数与仔猪窝均质量的对应关系

Fig. 9　Correspondence between average number of pixels per piglet and average weight of piglet litter

点数也在变化。由图 10 可知，平均每头仔猪像素点数与仔猪窝均质量呈正相关，通过计算可得仔猪平均每千克由 3 532 个像素点表示，仔猪窝均质量和平均每头仔猪像素点数之间皮尔逊相关系数为 0.956，使用 sklearn 库中线性回归模型进行回归预测，模型均方误差（MSE）为 0.042 kg^2。

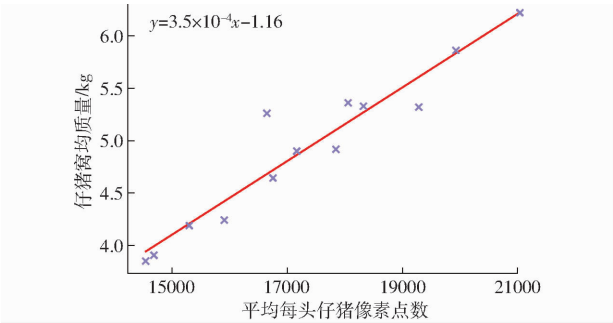

图 10　仔猪像素点数与仔猪窝均质量散点图

Fig. 10　Scatter plot of number of piglet pixels and average number of piglet nests

3.2.2　采样频率对结果稳定性的影响

在采样过程中，由于仔猪在产房中活动，导致仔猪像素点数在不断变化，采样频率过低会导致检测结果不稳定；在规模化猪场中，由于产房中栏数大，图像采样频率过高会导致硬件设备过载，因此针对采样频率对结果的稳定性影响进行实验分析。设置实验过程如下：首先设定采样频率，对 50 000 幅图像随机采样 1 000 次，每一次采样都会根据 OD_SeGAN 模型检测结果生成仔猪平均每千克对应的像素点数以及仔猪窝均质量和平均每头仔猪像素点数之间的皮尔逊相关系数两组数据，并计算两组数据的标准差；其次，改变采样频率，以标准差作为工具分析每次采样生成数据的稳定性。采样频率为 50 ~ 1 700 次/（8 h）。

图 11 为不同采样频率随机采样 1 000 次时，平均每千克对应的像素点数标准差以及仔猪窝均质量和仔猪窝均质量对应的像素点数之间的皮尔逊相关系数标准差，标准差越小表示该采样频率生成的数据越稳定。由图 11 可知，随着采样频率的增加，生成的相关数据趋于稳定。当采样频率为 1 000 次/（8 h），即从白天 8 h 采集的图像中随机采样1 000 幅，随机采样 1 000 次，平均每千克像素点数平均值为 3 531，标准差为 3.22 kg。仔猪窝均质量和仔猪窝均质量对应的像素点数之间的皮尔逊相关系数平均值为 0.956，标准差为 0.002 45，线性回归 MSE 为 0.043 kg^2。

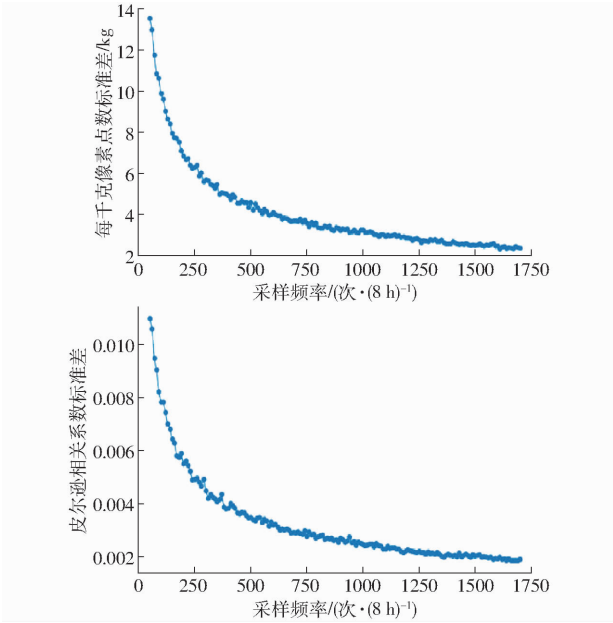

图 11　采样频率对结果稳定性的影响

Fig. 11　Effect of sampling frequency on stability of sampling results

4　结论

（1）提出一种基于 YOLO v5s 和 GAN 的仔猪实例分割模型 OD_SeGAN，采用 OD_SeGAN 模型分割产房中的仔猪，在测试集上 IoU 达 88.6%，高于 Mask_RCNN、Yolact 等实例分割模型。

（2）利用 OD_SeGAN 模型获取仔猪像素点数，使用平均每头仔猪像素点数衡量仔猪窝均质量，计算得仔猪窝均质量和平均每头仔猪像素点数之间皮尔逊相关系数为 0.956，线性回归 MSE 为 0.042 kg^2。

参　考　文　献

[1] NASIRAHMADI A, HENSEL O, EDWARDS S A, et al. Automatic detection of mounting behaviours among pigs using image analysis[J]. Computers and Electronics in Agriculture, 2016, 124: 295 - 302.

[2] 沈明霞，陈金鑫，丁奇安，等. 生猪自动化养殖装备与技术研究进展与展望[J]. 农业机械学报，2022，53(12)：1 - 19.
SHEN Mingxia, CHEN Jinxin, DING Qi'an, et al. Current situation and development trend of pig automated farming equipment

application[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 1 - 19. (in Chinese)
[3] 张站奇,孙显彬,孙贺,等. 基于改进 YOLO v5 - StrongSORT 的屠宰场猪只精准计数方法[J]. 农业机械学报,2024,55(12):354 - 364.
ZHANG Zhanqi, SUN Xianbin, SUN He, et al. Accurate counting of pigs in slaughterhouses based on improved YOLO v5 - StrongSORT[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(12):354 - 364. (in Chinese)
[4] 薛鸿翔,沈明霞,刘龙申,等. 基于改进 YOLO v5s 的经产母猪发情检测方法研究[J]. 农业机械学报,2023,54(1):263 - 270.
XUE Hongxiang, SHEN Mingxia, LIU Longshen, et al. Estrus detection method of parturient sows based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 263 - 270. (in Chinese)
[5] CANG Y, HE H, QIAO Y. An intelligent pig weights estimate method based on deep learning in sow stall environments[J]. IEEE Access, 2019, 7: 164867 - 164875.
[6] QIAO Y, KONG H, CLARK C, et al. Intelligent perception for cattle monitoring: a review for cattle identification, body condition score evaluation, and weight estimation[J]. Computers and Electronics in Agriculture, 2021, 185: 106143.
[7] 刘龙申,舒翠霓,李波,等. 基于 EfficientDet 的围产期母猪姿态识别[J]. 农业机械学报,2022,53(4):271 - 279.
LIU Longshen, SHU Cuini, LI Bo, et al. Recognition of perinatal sows' posture based on EfficientDet algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 271 - 279. (in Chinese)
[8] 李菊霞,李艳文,牛帆,等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报,2021,52(3):251 - 256.
LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251 - 256. (in Chinese)
[9] CHEN C, ZHU W, LIU D, et al. Detection of aggressive behaviours in pigs using a RealSence depth sensor[J]. Computers and Electronics in Agriculture, 2019, 166: 105003.
[10] 刘龙申,沈明霞,柏广宇,等. 基于机器视觉的母猪分娩检测方法研究[J]. 农业机械学报,2014,45(3):237 - 242.
LIU Longshen, SHEN Mingxia, BAI Guangyu, et al. Sows parturition detection method based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(3): 237 - 242. (in Chinese)
[11] 王荣,高荣华,李奇峰,等. 融合特征金字塔与可变形卷积的高密度群养猪计数方法[J]. 农业机械学报,2022,53(10):252 - 260.
WANG Rong, GAO Ronghua, LI Qifeng, et al. High-density pig herd counting method combined with feature pyramid and deformable convolution[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 252 - 260. (in Chinese)
[12] 赵宇亮,曾繁国,贾楠等. 基于 DeepLabCut 算法的猪只体尺快速测量方法研究[J]. 农业机械学报,2023,54(2):249 - 255,292.
ZHAO Yuliang, ZENG Fanguo, JIA Nan, et al. Rapid measurements of pig body size based on DeepLabCut algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 249 - 255,292. (in Chinese)
[13] CHEN L C, HERMANS A, PAPANDREOU G, et al. Masklab: instance segmentation by refining object detection with semantic and direction features[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4013 - 4022.
[14] 王海燕,江烨皓,黎煊,等. 基于弱监督数据集的猪只图像实例分割[J]. 农业机械学报,2023,54(10):255 - 265.
WANG Haiyan, JIANG Yehao, LI Xuan, et al. Pig image instance segmentation based on weakly supervised dataset[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(10):255 - 265. (in Chinese)
[15] LI Y, QI H, DAI J, et al. Fully convolutional instance-aware semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2359 - 2367.
[16] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 - 2969.
[17] YANG A, HUANG H, ZHENG C, et al. High-accuracy image segmentation for lactating sows using a fully convolutional network[J]. Biosystems Engineering, 2018, 176: 36 - 47.
[18] 高云,郭继亮,黎煊,等. 基于深度学习的群猪图像实例分割方法[J]. 农业机械学报,2019,50(4):179 - 187.
GAO Yun, GUO Jiliang, LI Xuan, et al. Instance-level segmentation method for group pig images based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 179 - 187. (in Chinese)
[19] HUANG E, HE Z, MAO A, et al. A semi-supervised generative adversarial network for amodal instance segmentation of piglets in farrowing pens[J]. Computers and Electronics in Agriculture, 2023, 209: 107839.
[20] 杨小娟,李世强,谢耀弟,等. 仔猪断奶个体重与保育期生长性能、腹泻率和死亡率的关系[J]. 甘肃畜牧兽医,2021,51(6):57 - 62.
[21] 李健源,柳春娜,卢晓春,等. 基于改进 YOLOv5s 和 TensorRT 部署的鱼道过鱼监测[J]. 农业机械学报,2022,53(12):314 - 322.
LI Jianyuan, LIU Chunna, LU Xiaochun, et al. Fish passage monitoring based on improved YOLOv5s and TensorRT deployment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 314 - 322. (in Chinese)
[22] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[J]. arXiv Preprint, arXiv:1511.07122, 2015.

[23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 – 7141.
[24] BOCHKOVSKIY, ALEXEY, WANG Chien-Yao, et al. Yolov4: optimal speed and accuracy of object detection[J]. arXiv Preprint, arXiv:2004. 10934,2020.
[25] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 – 13000.
[26] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146 – 157.
[27] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658 – 666.
[28] GEVORGYAN Z. SIoU loss: more powerful learning for bounding box regression[J]. arXiv Preprint, arXiv:2205. 12740, 2022.
[29] CAI Z, VASCONCELOS N. Cascade R – CNN: delving into high quality object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154 – 6162.
[30] BOLYA D, ZHOU C, XIAO F, et al. Yolact: real-time instance segmentation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9157 – 9166.
[31] WANG X, KONG T, SHEN C, et al. Solo: segmenting objects by locations[C]// Computer Vision-ECCV 2020: 16th European Conference, 2020: 649 – 665.

(上接第 **481** 页)

[22] YANG M, HAN G, YAN B, et al. Hybrid-sort: weak cues matter for online multi-object tracking[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2024: 6504 – 6512.
[23] 龚轩, 乐孜纯, 王慧, 等. 多目标跟踪中的数据关联技术综述[J]. 计算机科学, 2020, 47(10): 136 – 144
GONG Xuan, LE Zichun, WANG Hui, et al. Survey of data association technology in multi-target tracking[J]. Computer Science, 2020, 47(10): 136 – 144. (in Chinese)
[24] SUN P, CAO J, JIANG Y, et al. Dancetrack: multi-object tracking in uniform appearance and diverse motion[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 20993 – 21002.
[25] SUN Z, CHEN J, CHAO L, et al. A survey of multiple pedestrian tracking based on tracking-by-detection framework[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(5): 1819 – 1833.
[26] JOCHER G, CHAURASIA A, QIU J. Ultralytics YOLOv8[Z]. 2023. https://github. com/ultralytics/ultralytics.
[27] LIU Z, WANG X, WANG C, et al. Sparsetrack: multi-object tracking by performing scene decomposition based on pseudo-depth[J]. arXiv Preprint, arXiv:2306. 05238, 2023.
[28] BERNARDIN K, STIEFELHAGEN R. Evaluating multiple object tracking performance: the clear mot metrics[J]. EURASIP Journal on Image and Video Processing, 2008,2008: 1 – 10.
[29] RISTANI E, SOLERA F, ZOU R, et al. Performance measures and a data set for multi-target, multi-camera tracking[C]// European Conference on Computer Vision, 2016: 17 – 35.
[30] LIN T-Y, MAIRE M, BELONGIE S, et al. Microsoft coco: common objects in context[C]// Computer Vision-ECCV 2014: 13th European Conference, 2014: 740 – 755.