

doi:10.6041/j.issn.1000-1298.2020.04.029

基于改进 YOLO v3 模型的挤奶奶牛个体识别方法

何东健^{1,2} 刘建敏^{1,2} 熊虹婷^{1,3} 芦忠忠^{1,2}

- (1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;
2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100;
3. 陕西省农业信息感知与智能服务重点实验室, 陕西杨凌 712100)

摘要:为实现无接触、高精度养殖场环境下奶牛个体的有效识别,提出了基于改进 YOLO v3 深度卷积神经网络的挤奶奶牛个体识别方法。首先,在奶牛进、出挤奶间的通道上方安装摄像机,定时、自动获取奶牛背部视频,并用视频帧分解技术得到牛背部图像;用双边滤波法去除图像噪声,并用像素线性变换法增强图像亮度和对比度,通过人工标注标记奶牛个体编号;为适应复杂环境下的奶牛识别,借鉴 Gaussian YOLO v3 算法构建了优化锚点框和改进网络结构的 YOLO v3 识别模型。从 89 头奶牛的 36 790 幅背部图像中,随机选取 22 074 幅为训练集,其余图像为验证集和测试集。识别结果表明,改进 YOLO v3 模型的识别准确率为 95.91%,召回率为 95.32%,mAP 为 95.16%,IoU 为 85.28%,平均帧率为 32 f/s,识别准确率比 YOLO v3 高 0.94 个百分点,比 Faster R-CNN 高 1.90 个百分点,检测速度是 Faster R-CNN 的 8 倍,背部为纯黑色奶牛的 F1 值比 YOLO v3 提高了 2.75 个百分点。本文方法具有成本低、性能优良的特点,可用于养殖场复杂环境下挤奶奶牛个体的实时识别。

关键词: 挤奶奶牛; 个体识别; 视频分析; 机器视觉; 改进 YOLO v3

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2020)04-0250-11 OSID: 

Individual Identification of Dairy Cows Based on Improved YOLO v3

HE Dongjian^{1,2} LIU Jianmin^{1,2} XIONG Hongting^{1,3} LU Zhongzhong^{1,2}

- (1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
3. Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China)

Abstract: Aiming to achieve an effective identification of dairy cows in a non-contact and high-precision environment of farming, a method to identify dairy cows based on the improved YOLO v3 deep convolutional neural network was proposed. According to this method, multiple cameras were installed above the passageway between the doors of the milking room. The back of cows was videotaped automatically and regularly, after which the image of the cow's back was captured by applying video frame decomposition technology. Upon the removal of images noise with bilateral filters and the enhancement of brightness and contrast with the pixel linear transformation method, the individual dairy cows were serial numbered manually. For the cows to be better identified in complex environments, the YOLO v3 recognition model that features optimized anchor boxes and improved network structure was constructed by making reference to the Gaussian YOLO v3 algorithm. From totally 36 790 images showing the back of 89 cows, 22 074 were randomly selected as the training set, while the remaining ones were classified into either the validation set or the test set. The results showed that the accuracy of the improved YOLO v3 was 95.91%, the recall rate was 95.32%, the mAP was 95.16%, the IoU was 85.28%, the actual frame rate of detection was 32 f/s, and the accuracy rate of identification was 0.94 percentage points higher compared with that of the YOLO v3 and 1.90 percentage points higher than that of Faster R-CNN. Moreover, the detection speed was eight times faster than that of Faster R-CNN, while the F1 value of dairy cows with pure black back was 2.75 percentage points higher compared with that of the original algorithm. The method showed such advantages as low cost and excellent performance, which

were not only conducive to the real-time identification of dairy cows in complex farm environments, but also to the extended application of this method to the identification of other large-sized animals.

Key words: dairy cow; individual identification; video analysis; machine vision; improved YOLO v3

0 引言

以信息与智能技术为支撑的绿色、高效、精准养殖成为现代畜牧业发展的必然趋势^[1-3]。奶牛养殖业作为中国畜牧业的重要组成部分,其信息化、自动化、智能化应用不断深化^[4-7]。精准养殖体系中个体档案建立、信息采集、执行方案制定以及产品溯源等,都需要对奶牛进行快速准确的识别^[2]。个体身份识别是对奶牛个体进行管理以及自动分析奶牛行为的前提和应用基础^[8],对奶业发展具有至关重要的作用。对于现代化的规模化奶牛养殖场,挤奶厅不仅完成挤奶作业,更是养殖场数据采集的重要场所,在这里可以监测每头牛的产奶量、挤奶时间、日挤奶次数、产奶品质等数据。对进入挤奶厅的每头奶牛进行准确识别,可以在未来与不同牛只个体的产奶情况准确匹配,这对企业掌握每头奶牛的产奶量情况、提高并改进对奶牛的喂养管理、判断奶牛健康状况、增加养殖场经济效益等具有重要意义。传统的奶牛个体识别主要采用人工耳标、信息登记、佩戴颈链等方法^[9],依靠人工观察方式不能适应现代化规模养殖自动识别的应用需求。无线射频识别技术(Radio frequency identification, RFID)在实时更新资料、存储信息量、使用寿命、工作效率、安全性等方面都具有优势,近年来在畜牧业中被用于动物识别^[10-11]。典型的电子识别装置必须通过穿刺固定在颈部或耳朵上,这会对奶牛造成损害并影响动物福利,且 RFID 电子标签价格较高,在规模化养殖中需要多台 RFID 读取设备识别奶牛,大大增加了生产成本。

随着计算机技术的发展,结合荷斯坦奶牛特有的黑白花纹信息,采用计算机视觉实现奶牛个体识别逐步引起国内外学者的关注^[12-17]。LI 等^[18]提出自动识别奶牛尾部图像作为感兴趣区域(Region of interest, ROI),以 Zernike 矩描述区域内白色图案的形状特征,采用支持向量机(Support vector machine, SVM)对 23 头奶牛分类的最高精度达到 99.7%,但对 ROI 为纯黑色的奶牛并不适用。CAI 等^[19]将牛脸图像分割为互不重叠的独立小块,使用局部二值模式方法提取纹理特征,并设计了特征描述子,建立的面部描述模型能准确高效地进行牛脸识别,但实际场景下牛的头部不断移动,难以自动采集其脸部图像。张满囤等^[20]提出一种基于小波变换和改进

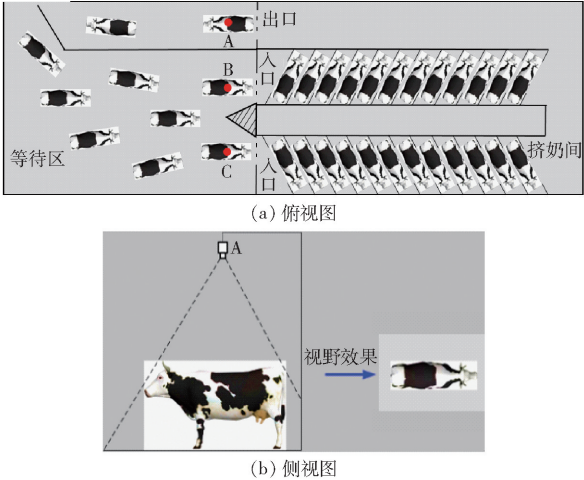
核主成分分析法的奶牛个体识别方法,但该方法实时性较差。赵凯旋等^[21]提出一种基于卷积神经网络的奶牛个体识别方法,该方法对 30 头奶牛的识别正确率为 93.33%,但训练样本数较少,不能满足中等规模养殖场的需求。ANDREW 等^[22]研究采用区域卷积神经网络(RCNN)对荷斯坦黑白花牛个体识别的方法,在 89 头奶牛中准确率达到 86.1%,对无人机拍摄的 23 头室外放牧奶牛的识别准确率达 98.1%,但该研究针对室外放牧环境,不适合挤奶厅奶牛的识别。YOLO(You only look once)算法^[23-25]是一种端到端的卷积神经网络,检测速度快,在农业目标实时检测与识别中具有良好应用前景^[26]。深度学习在农业对象的识别中表现出良好的性能和抗干扰能力,具有复杂场景下奶牛个体识别的潜力。

为提供一种可在实际奶牛养殖环境下应用、非接触、低成本的挤奶奶牛个体识别方法,本文在设计奶牛进出挤奶厅背部视频采集方案的基础上,提出基于视频分析和改进 YOLO v3 深度卷积神经网络的挤奶奶牛个体识别方法。

1 供试图像采集与预处理

1.1 图像采集

本研究供试视频拍摄于陕西省杨凌科元克隆股份有限公司,拍摄对象为 89 头泌乳期间健康的美国荷斯坦奶牛。每天早、中、晚进行 3 次挤奶,在挤奶时间段内,奶牛从等待区的 2 个入口进入挤奶厅,每个挤奶厅有 12 台挤奶设备,完成挤奶的奶牛从唯一出口离开。如图 1 所示,用 A、B、C 表示的 3 台 DS-2CD3325-I 型高清网络摄像机(海康威视,杭州)分



(a) 俯视图

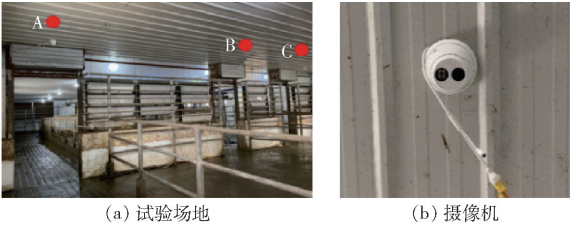
(b) 侧视图

图 1 摄像机安装位置示意图

Fig. 1 Installation location of video camera

别安装在出入口的正上方,距离地面高度为 2.7 m,调整摄像机视轴竖直向下,并使其视场水平方向与牛只行走方向基本平行,视野宽度大于 1.5 个牛身长度。

每个摄像机有一个独立的 IP 地址,通过摄像机配套的萤石云客户端,可远程实时访问摄像机,以浏览对应位置的奶牛背部视频。同时,摄像机采集的视频通过网络传输至 25 m 外挤奶厅监控室的硬盘录像机内,录像机接入养殖场内部有线网络接口。视频采集系统如图 2 所示。



(a) 试验场地 (b) 摄像机

图 2 视频采集系统

Fig. 2 Video acquisition system

于 2019 年 3 月 17—21 日拍摄供试视频,根据养殖场的固定挤奶时间设置摄像机拍摄时间段,每天 05:00—07:00、13:00—15:00、19:00—21:00 时段自动获取奶牛进出挤奶厅时的背部视频片段。89 头奶牛按 12 头 1 组随机从不同入口进入挤奶间(最后一组为 5 头)。当每组第 1 头奶牛全部出现在视野左侧时开始采集,到最后一头奶牛行走至视野右侧边缘结束,以此作为一个视频段。

对不同时段进出挤奶厅的每组奶牛背部视频,剔除包含奶牛停顿和异常行为的视频后,共采集到 48 段视频,每段视频时长在 20 ~ 80 s 之间,视频帧率为 25 f/s,分辨率为 1 920 像素(水平) × 1 080 像素(垂直)。由于夜间挤奶厅和等待区有灯光照明,故夜间拍摄的视频段也被选用。利用视频帧分解技术,每 3 帧取 1 帧,得到无异常的奶牛背部图像 10 800 幅。

用于图像数据处理及模型训练的计算机处理器为 Inter Core i5 - 8700,主频为 3.20 GHz,显卡为 NVIDIA GeForce GTX 1070 Ti,16 GB 内存,500 GB 硬盘。所有程序均在 Windows 10 系统下用 Python 语言编写,图像显示处理调用 OpenCV 库、数据计算调用 CUDA、CuDNN 等。

1.2 视频图像预处理

由于奶牛养殖场环境复杂、光照不均等,导致视频帧图像中包含噪声及边缘模糊等,需要去除视频帧图像中的噪声,并做图像增强等预处理,以提高有效信息占比,为后续目标检测、特征提取和个体识别等奠定良好的基础。

1.2.1 双边滤波去噪

实际拍摄的视频地面颜色与奶牛背部黑色斑点颜色相近,若采用高斯滤波或中值滤波,在边缘处易出现模糊现象,考虑到双边滤波算法^[27]结合图像的空间邻近度和像素值相似度进行折中处理,能在保持强边缘的同时有效地对图像的细小变化进行平滑,故用双边滤波算法去除噪声。

为验证双边滤波算法对本文图像去噪的有效性,用双边滤波算法、高斯滤波、中值滤波、均值滤波、方框滤波 5 种方法对同一幅图像进行去噪效果对比试验,不同滤波器处理后图像的峰值信噪比(PSNR)和结构相似性值(SSIM)如表 1 所示。

表 1 不同滤波器下图像的峰值信噪比和结构相似性值

Tab. 1 PSNR and SSIM of image under different filters

图像处理方法	峰值信噪比/dB	结构相似性值
双边滤波	36.23	0.999 98
高斯滤波	30.57	0.999 77
中值滤波	28.72	0.999 90
均值滤波	27.09	0.995 14
方框滤波	27.09	0.995 14

由表 1 可知,原始图像经过双边滤波算法处理后,峰值信噪比为 36.23 dB,与原始图像结构相似性值达到 0.999 98,在 5 种滤波器中均取得最高值,图像质量相对较好。故本文选择双边滤波算法去除噪声。

1.2.2 图像增强

由于挤奶厅地面为深灰色,与奶牛身体颜色相近,使得图像偏暗,且奶牛躯干边缘轮廓和地面难以明显区分,故需对图像像素进行亮度、对比度增强,增强公式为

h(x)=αf(x)+β (1)

式中 h(x)——输出图像 f(x)——输入图像
α——增益,影响图像对比度,预备试验确定 α 为 1.1

β——偏置,影响图像亮度,预备试验确定 β 为 30

图像增强后,通过自定义的卷积核 k_{ernel} 进行锐化处理,即

k_{kernel} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} (2)

通过强化奶牛躯干边缘轮廓与背部花斑分界线,达到增强边缘的效果。

1.3 摄像机标定

DS-2CD3325-I 型半球形摄像机为广角摄像机(焦距为 2.8 mm,水平视场角为 103°),图像呈现

严重的桶形失真^[28],如图3a所示。

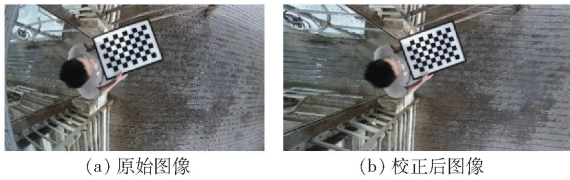


图3 畸变校正结果

Fig. 3 Result of distortion correction

为了消除图像畸变,本文采用张正友标定方法进行标定。采用边长为0.5 m×0.35 m、印制有间隔0.05 m黑白网格的棋盘格标定板,获取视场中与奶牛同一高度的不同位置、不同角度下的25幅图像,利用Matlab中CameraCalibrator标定工具箱对每一幅图像提取角点信息,依次连接各个内角点,自动获取摄像机畸变系数和内参数矩阵。相机坐标系(x, y, z)与世界坐标系(X, Y, Z)的对应关系可表示为

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3)$$

式中 f_x —— x 轴焦距 f_y —— y 轴焦距
 c_x —— x 轴光学中心
 c_y —— y 轴光学中心

可在Python中查看标定结果,实现图像畸变校正。

校正后根据有效区域进行剪裁,最终得到的奶牛背部图像尺寸为1 444像素(水平)×756像素(垂直)。矩形网格畸变校正结果如图3b所示。

1.4 供试数据

为避免样本量差异对识别结果的影响,需要保证不同模式中参与训练的样本数量基本一致,在经过预处理的10 800幅图像中,选择7 358幅包含独立完整目标的奶牛背部图像。结合奶牛进入挤奶厅时躯干与视野水平线有一定倾斜角度、视频图像中光照强度随时间变化的特点,对原始数据集采用旋转10°、旋转-10°、亮度线性增强10、亮度线性降低10的方法进行数据扩充,最终得到36 790幅图像。随机选取其中22 074幅作为训练集,7 358幅为验证集,7 358幅为测试集,数据之间无重叠。

为了提供训练模型的数据类别,采用开源工具Labellmg对训练和测试图像集进行人工标注。由于奶牛行走中头部会产生起伏,有抬头和低头行为,使得奶牛头部形状变化较大,且头部占奶牛身体的比例较小,故标注时忽略奶牛的头部,仅框选奶牛背部区域,将其标记为对应奶牛的身份ID(耳标),并自动生成相应的配置文件。

2 YOLO v3 识别模型的构建与改进

深度学习具有强大的数据表征能力,通过组合

低层特征形成更抽象的高层表示属性类别或特征,在样本量足够大时可以达到较好的识别精度^[29],适用于本文奶牛个体的识别。当前以R-CNN系列^[30-31]为代表的基于区域分析的模型和以SSD^[32]、YOLO系列为代表的基于回归分析的模型应用较为广泛。其中,YOLO v3识别模型具有轻量高速的特点,在快速检测的同时识别准确率高,非常适合作为奶牛实时识别模型。

2.1 边界框的分类和预测

YOLO v3将输入图像首先缩放至416像素×416像素,然后划分为 $S \times S$ 个网格,输入神经网络进行训练。每个网格单元负责检查其内的边界框和其置信度,网格单元具体信息可用 $T(x, y, w, h, C)$ 表示, x, y 是当前网格单元预测检测对象的置信度中心位置的横、纵坐标, w, h 是外接矩形的宽度和高度,置信度 C 则反映当前网格单元是否包含奶牛目标及其预测准确性,通过阈值对预测结果进行取舍。假设图像的左上角为坐标原点,则预测的边界框可表示为

$$\begin{cases} b_x = \sigma(x) + d_x \\ b_y = \sigma(y) + d_y \\ b_w = p_w e^w \\ b_h = p_h e^h \end{cases} \quad (4)$$

式中 b_x, b_y ——预测边界框中心点的横、纵坐标值
 $\sigma(x), \sigma(y)$ ——预测边界框中心点与最近网格边缘在 x, y 向的距离

d_x, d_y ——网格的横、纵坐标偏移量

b_w, b_h ——预测边界框的宽度和高度

p_w, p_h ——锚点框的宽度和高度

位置边框采用logistic预测目标对象的置信度。如果预测边界框与真实边界框重叠,并且预测结果优于所有其他边界,则该框的值为1,否则为0。

2.2 YOLO v3 模型的改进

本文识别的荷斯坦奶牛身形较大,且不同个体之间体长、体宽相近,YOLO v3模型的先验锚点框和结构并不适于奶牛识别。故本文使用K-means算法对数据集中的奶牛目标聚类分析,针对奶牛识别修改模型的层级结构,并参考Gaussian YOLO v3^[33]构建改进YOLO v3模型。

2.2.1 锚点框的聚类与优化

YOLO v3借鉴了Faster R-CNN和SSD中锚点框作为初始候选框的方法,其大小为人工设定的固定值,在不同尺度上对目标进行预测,初始锚点框的选择影响网络的速度和精度。为加快网络收敛速度,本文用K-means算法对数据集中标注的边界框

进行聚类,得到适合奶牛数据集的锚点框。为了减少锚点框尺寸对误差的影响,用平均交并比 (Average intersection over union, AIoU) 代替欧氏距离作为 K-means 聚类分析的评判标准,平均交并比越大表示聚类效果越好,距离表示为

$$D_{A,T} = \sum_{i=1}^K \sum_{j=1}^{N_i} (1 - I_{oU}) \tag{5}$$

式中 A——真实的边界框区域
T——聚类得到的矩形框区域
 $D_{A,R}$ ——真实框中心与聚类中心之间的距离
K——聚类中心(即锚点框)数量
 N_i ——第 i 个聚类中心的样本数量
 I_{oU} ——真实框和聚类框的交并比

令 $K=1,2,\cdots,12$ 分别对数据集样本进行聚类分析,得到平均交并比与 K 之间的关系如图 4 所示。

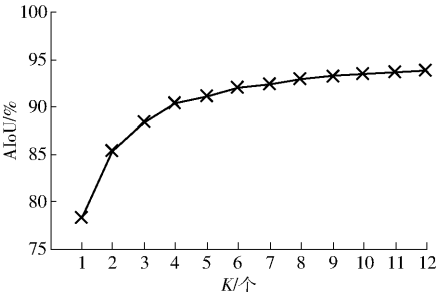


图 4 平均交并比与 K 之间的关系

Fig. 4 Relationship between AIoU and K

由图 4 可知,当锚点框数量为 9 时,平均交并比达到 93.5%,且此后变化平稳。为平衡交并比与网络复杂度,将 $K=9$ 的聚类结果作为网络中的锚点框尺寸,即 (199, 165)、(208, 190)、(229, 177)、(233, 201)、(222, 222)、(260, 196)、(272, 208)、(255, 234)、(235, 274)。

2.2.2 网络结构的改进

YOLO v3 采用 Darknet-53 网络结构,在前向传播过程中,通过改变卷积核步长实现张量的尺寸变化。Darknet-53 借鉴了残差神经网络^[34]的构造,每个残差模块由多个残差单元(res unit)组成,每个残差单元有 2 个卷积层和 1 个快捷链路,使网络结构加深的同时避免梯度消失,加强对图像特征的学习。输入通过 2 个 DBL (Darknetconv2d - BN - Leaky) 单元进行残差操作,其中,DBL 单元包含卷积层、批归一化层(BN)和 Leaky ReLU 激活函数层,残差单元的基本结构如图 5 所示。

YOLO v3 模型(不含图 6 中虚线框包围部分)中有 5 次下采样,输出 3 个不同尺寸的特征图。将更深层网络的特征进行步长为 2 的上采样,与前一步下采样得到的特征进行张量拼接,使网络既可以

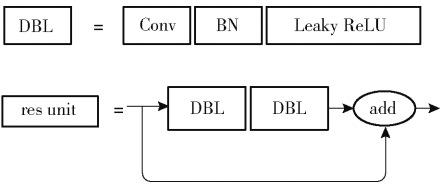


图 5 残差单元的基本结构

Fig. 5 Structure of res unit

包含高层特征的高级语义信息又可以保留低层特征的物体位置信息。网络在 13×13 (尺度 1)、 26×26 (尺度 2)、 52×52 (尺度 3) 3 种尺度上对不同大小的目标进行边界框预测,在 COCO 数据集中对应的锚点框是 (10, 13)、(16, 30)、(33, 23)、(30, 61)、(62, 45)、(59, 119)、(116, 90)、(156, 198)、(373, 326),而奶牛数据集锚点框最小值(199, 165)大于前 8 个锚点框,因此训练时只能在 1 个尺度上提取特征,特征提取效率低。为提高网络收敛速度,增强特征提取能力,故将本文经 K-means 聚类得到的锚点框按从大到小的顺序分配给尺度 1、尺度 2 和尺度 3,每个尺度预测 3 个锚点框,以提高多尺度目标预测准确率。

本文识别奶牛数量为 89 个,样本数量和计算量较大,因此在残差模块和降采样层间增加 1 个 1×1 卷积层,在不损失分辨率的前提下进一步降低输入通道数、减少计算量。改进后的网络层数更深,为防止随着网络层数的增加可能造成的梯度消失,提升网络训练效果,将 YOLO v3 输出模块中的前 4 个卷积层替换为 2 个残差单元,通过跳跃连接的方式将信息传递到更深层。改进内容用虚线框标出,则改进后的网络结构如图 6 所示。

2.2.3 改进 YOLO v3 模型的构建

YOLO v3 模型输出预测框的位置参数 b_x, b_y, b_w, b_h 和目标置信度,边界框和置信度均单独回归。输出的预测框只有确定性坐标值而无概率值,无法表示当前预测框的准确性。为进一步提高检测精度,本文借鉴 Gaussian YOLO v3^[33],采用高斯模型对预测框的可靠性进行建模。给定输入 x_g 、输出 y_g 的单个高斯模型如下

$$G(y_g | x_g) = N_g(y_g | \mu(x_g), \Sigma(x_g)) \tag{6}$$

式中 $\mu(x_g)$ ——均值,表示预测边框的位置
 $\Sigma(x_g)$ ——方差,表示预测框位置的不确定性
 G, N_g ——高斯分布参数

将预测特征图中每个预测框的坐标用均值和方差建模,预测框的输出更改为 $\mu(b_x), \Sigma(b_x), \mu(b_y), \Sigma(b_y), \mu(b_w), \Sigma(b_w), \mu(b_h), \Sigma(b_h)$,通过 sigmoid 函数将输出值限定在 0 ~ 1 之间。

2.2.4 损失函数

YOLO v3 使用方差相加与二值交叉熵的损失计

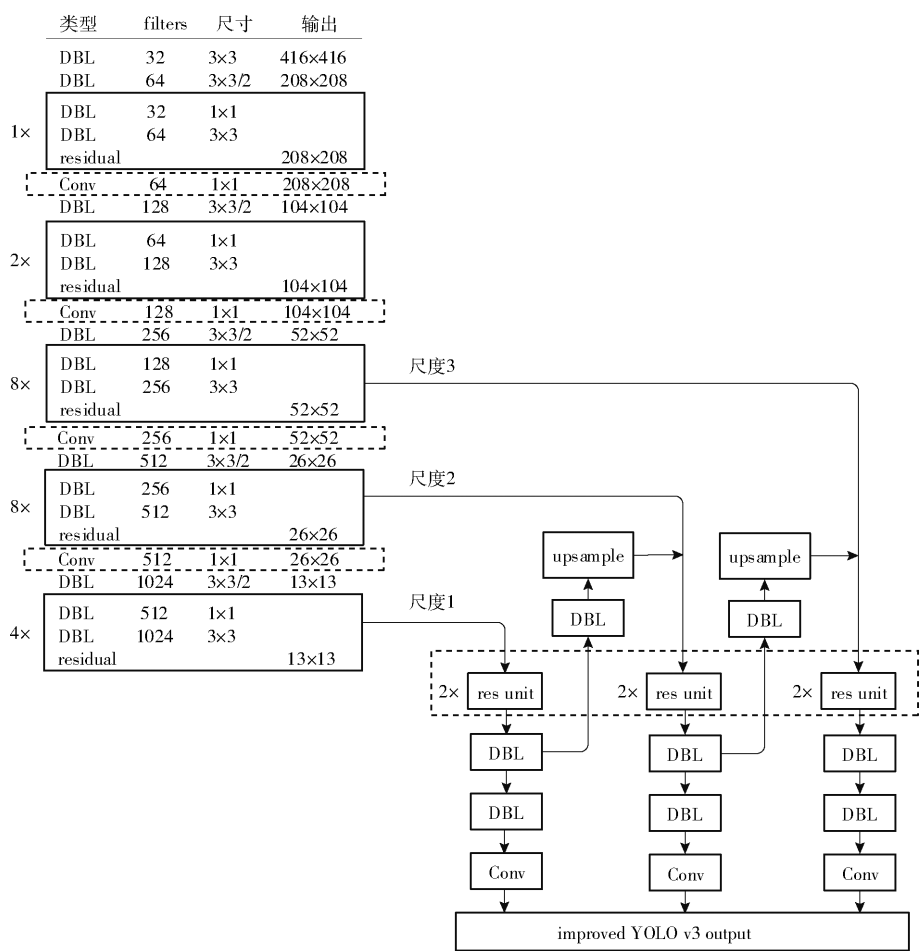


图 6 改进后的 YOLO v3 网络结构

Fig.6 Structure of improved YOLO v3

算方法,损失函数由坐标误差、交并比(Intersection over union,IoU)误差和分类误差 3 部分组成^[23-25]。在训练时通过反向传播不断更新模型,使得损失值不断减小。通过高斯建模将预测框坐标输出为高斯参数,因此预测框的损失函数更改为负对数似然损失^[33],即

$$L_x = - \sum_{i=1}^W \sum_{j=1}^H \sum_{k=1}^K \gamma_{ijk} \lg(N(x_{ijk}^G | \mu_{b_x}(x_{ijk}), \Sigma_{b_x}(x_{ijk}))) + \varepsilon) \tag{7}$$

式中 L_x ——坐标 b_x 的负对数似然损失,与 L_y 、 L_w 、 L_h 的计算方法相同
 W 、 H ——高度和宽度的网格数
 γ_{ijk} ——对数似然损失参数
 x_{ijk}^G —— b_x 的真值
 x_{ijk} ——网格 (i,j) 中的第 k 个锚点
 $\mu_{b_x}(x_{ijk})$ —— b_x 的均值坐标
 $\Sigma_{b_x}(x_{ijk})$ —— b_x 的不确定性
 ε ——对数函数的数值稳定性参数

2.3 基于改进 YOLO v3 的奶牛个体识别模型训练

用 22 074 幅训练集图像进行训练,用其余图像进行验证和测试。将经畸变矫正后尺寸为 1 444 像

素 × 756 像素的图像转换为 416 像素 × 416 像素,以适于 YOLO v3 的输入。然后再分成 13 像素 × 13 像素的网格单元,以便于输入网络进行训练。训练时,以 64 幅图像为一个批次,每训练一批图像,更新一次权值参数。根据预备试验结果,权值的衰减速率(decay)设为 0.000 5,动量因子(momentum)设为 0.9,最大训练次数设置为 80 000,初始学习率(learning rate)设为 0.000 5。在迭代次数为 60 000 和 70 000 时,学习率降低为初值的 10% 和 1%,使模型在训练后期振荡减小,从而更加接近最优解。

本文对改进 YOLO v3 模型进行 80 000 次迭代,其损失(Loss)函数的变化如图 7 所示。由图 7 可知,在前 2 000 次的迭代中损失值迅速下降,表明模型快速拟合;在 2 000 ~ 30 000 次迭代中损失值缓慢减小,而 30 000 次迭代后损失值稳定在 0.1 ~ 0.2 之间,只有轻微振荡,表明此时的训练结果较好。

为了防止因迭代次数过多而产生过拟合,在 30 000 次迭代后,每隔 2 000 次迭代输出一次权值模型,得到 25 个模型。为了选取合适的模型,采用准确率(P)、召回率(R)、平均精度(mAP)、交并比

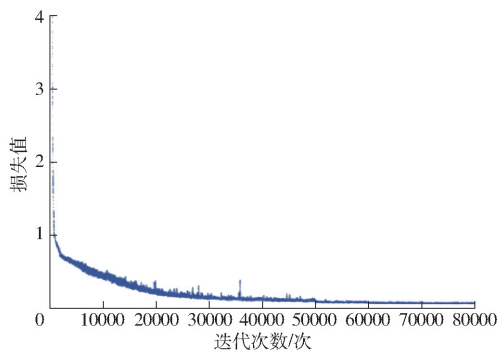


图 7 损失函数的输出结果

Fig. 7 Output result of loss function

(IoU) 和平均帧率(fps) 来评估模型性能, 计算公式如下

$$\begin{cases} P = \frac{T_p}{T_p + F_p} \times 100\% \\ R = \frac{T_p}{T_p + F_N} \times 100\% \\ m_{AP} = \frac{1}{N_{cow}} \sum_{n=1}^{N_{cow}} P(n) R(n) \times 100\% \\ I_{oU}(A, A') = \frac{|A \cap A'|}{|A \cup A'|} \times 100\% \end{cases} \quad (8)$$

式中 T_p ——真实正样本数量
 F_p ——虚假正样本数量
 F_N ——虚假负样本数量
 N_{cow} ——所识别奶牛个体类别数量
 n ——类别序号
 $P(n)$ ——第 n 类的准确率
 $R(n)$ ——第 n 类的召回率
 A' ——检测出的边界框区域

本文需要识别 89 头奶牛个体, 因此 N_{cow} 为 89。在性能评估参数中, IoU 用于测量真实值和预测值之间的相关度, 相关度越高, 该值越高, IoU 的阈值选择直接影响准确率和召回率。本文的目的是识别出相似背景下不同牛只个体编号, 对交并比要求并不高, 故以 IoU 是否大于 50% 判断奶牛目标是否被正确识别。

同时引入衡量分类模型精确度的 F1 指标, F1 同时兼顾了分类模型的准确率和召回率, 计算公式为

$$F_1 = \frac{2PR}{P + R} \quad (9)$$

式中 F_1 ——指标 F1

为了对模型的性能进行恰当的排序, 需要明确性能参数的优先级。在奶牛个体识别系统中, 试验采用的性能参数优先级由大到小依次为 mAP、F1、 P 、 R 、fps、IoU。

3 结果与讨论

mAP 是评价模型性能的主要指标, 故选择 mAP 最高的模型作为最优模型。通过最优模型对测试集中的奶牛个体进行识别, 得到 P 、 R 、F1、IoU、fps, 并与 YOLO v3、Faster R-CNN 进行比较。

3.1 最优模型的确定

改进 YOLO v3 模型迭代 30 000 次后输出的 mAP 变化曲线如图 8 所示。由图 8 可知, 当迭代次数为 70 000 次时, mAP 达到最大值 95.16%, 故选择迭代 70 000 次的权重参数作为最优模型参数。

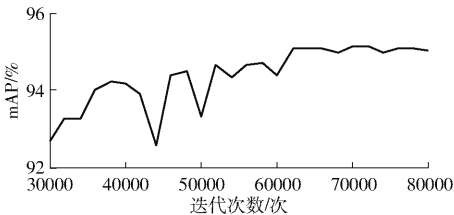


图 8 mAP 随迭代次数的变化曲线

Fig. 8 Change curve of mAP with iterations

置信度反映当前模型检测预测框内存在目标的可能性和目标位置的正确性, 预先设置置信度阈值, 预测框的置信度大于该值时被检测到。因此, 同一模型在不同置信度阈值下预测的 F1、 P 、 R 、IoU 会产生不同变化, 为找出最优置信度阈值, 图 9 给出不同阈值下的参数值。

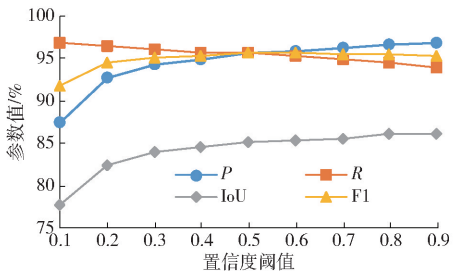


图 9 不同置信度阈值下的性能参数值的变化

Fig. 9 Changes of different parameters with different thresholds

由图 9 可知, 当置信度阈值取 0.6 时, F1 达到最大值 95.61%, 其他参数处于较高值, 模型预测结果最优。

3.2 模型识别结果

为了验证本文方法的可靠性和稳定性, 用测试集图像进行分类识别试验, 并与 YOLO v3、Faster R-CNN 识别结果进行对比。奶牛识别结果示例如表 2 所示。由表 2 可知, 3 种模型均能正确识别出不同奶牛个体编号, 当视野中有多头奶牛时也能被很好地识别出来, 表明构建的网络能可靠地识别奶牛, 对牛只个体的形变具有较强的泛化能力。

3 种模型在测试集上的 F1、准确率、召回率、mAP、IoU 和平均帧率如表 3 所示。由表 3 可知,本文模型测试集识别准确率为 95.91%,略高于 YOLO v3 的 94.97% 和 Faster R-CNN 的 94.01%;召回率为 95.32%,较 YOLO v3 提高了 4.98 个百分点,较 Faster R-CNN 提高了 6.03 个百分点;平均精度为 95.16%,比 YOLO v3 和 Faster R-CNN 分别提高了

2.36 个百分点和 2.41 个百分点。IoU 为 85.28%,比 Faster R-CNN 低 2.1 个百分点,但本文目的是识别奶牛个体,对奶牛位置精确度无要求,故 IoU 不影响对奶牛的识别。

从处理速度上看,本文模型平均帧率达 32 f/s,是 Faster R-CNN 的 8 倍,能够满足奶牛养殖场对进出挤奶厅的奶牛个体编号的实时识别要求。

表 2 奶牛识别结果示例

Tab. 2 Examples of identification results for dairy cows

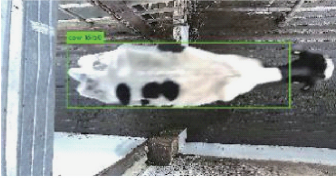
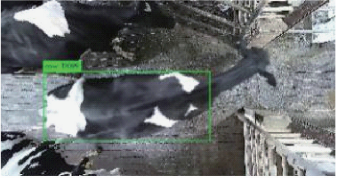
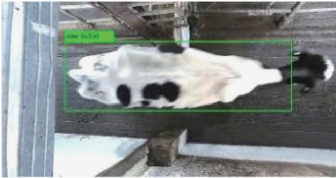
模型	16150号奶牛	12099号奶牛	08118号和16096号奶牛
本文模型			
YOLO v3			
Faster R-CNN			

表 3 3 种模型对奶牛个体的识别结果

Tab. 3 Recognition results of three models for dairy cows

模型	F1/%	准确率/%	召回率/%	mAP/%	IoU/%	平均帧率/(f·s ⁻¹)
本文模型	95.61	95.91	95.32	95.16	85.28	32
YOLO v3	92.60	94.97	90.34	92.80	82.51	38
Faster R-CNN	91.59	94.01	89.29	92.75	87.38	4

3.3 不同体斑颜色奶牛的对比试验

由于荷斯坦奶牛的身体斑点仅包含黑白两种颜色,因此,奶牛背部体斑的颜色是特征提取和识别的主要因素。当奶牛体斑颜色接近纯黑色时,识别难度大。供试 89 头泌乳期荷斯坦奶牛中,背部为纯黑色的奶牛约占 10%。

为了验证本文模型对不同体斑颜色的识别精度,构建体斑颜色测试集,共 300 幅图像,其中背部为纯黑色的单头奶牛图像 128 幅,背部为黑白双色的单头奶牛图像 172 幅。本文模型与 YOLO v3、Faster R-CNN 模型对不同体斑颜色奶牛检测的 F1 如表 4 所示。

由表 4 可知,本文模型对不同体斑颜色奶牛识别的 F1 值均高于 YOLO v3 和 Faster R-CNN 模型 2 个百分点左右。3 种模型对黑白双色奶牛识别的 F1 值高于对纯黑色奶牛的检测结果,这是因为黑色

表 4 3 种模型对不同体斑颜色奶牛的识别结果

Tab. 4 Recognition results of three models under different body spot characteristics of dairy cows

奶牛背部体斑颜色	样本数量/幅	模型	F1/%
黑白双色	172	本文模型	95.72
		YOLO v3	93.12
		Faster R-CNN	92.46
纯黑色	128	本文模型	91.53
		YOLO v3	88.78
		Faster R-CNN	89.24

奶牛本身特征参数较少,其特征提取更为困难。

3.4 不同模型的对比试验

为了验证本文模型的有效性,进行与其他模型的对比试验,结果如表 5 所示。

由表 5 可知,在识别对象和样本数量相当的情况下,本文模型的准确率比文献[15]模型提高了

表 5 不同模型识别结果

Tab.5 Recognition results of different models

模型	牛只数量	准确率/%	单帧识别时间/s
本文模型	89	95.91	0.031
YOLO v3	89	94.97	0.026
Faster R-CNN	89	94.01	0.253
文献[15]模型	16	84.20	
文献[16]模型	15	95.30	1.941
文献[22]模型	89	86.10	

11.71 个百分点,比文献[16]模型提高了 0.61 个百分点,单帧识别时间减少了 1.910 s,比文献[22]模型准确率提高了 9.81 个百分点。

上述结果表明,本文模型提高了奶牛个体识别准确率,实际检测平均帧率达 32 f/s,为后续实时检测奶牛挤奶时间和产奶量提供了技术支持。

3.5 讨论

本文构建的识别模型准确率为 95.91%,有部分奶牛未能正确识别。其中一头黑色奶牛错误识别的情况如图 10 所示,16054 号奶牛被误识别为 16079 号奶牛。

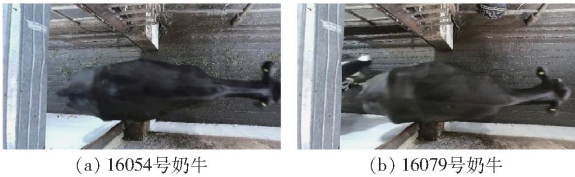


图 10 具有相似颜色的奶牛错误识别示例

Fig.10 Example of wrong identification of dairy cows with similar color

导致误识别的原因是 16054 号和 16079 号奶牛身体区域均接近纯黑色,背部花斑颜色相似,缺少用于识别的特征,即使肉眼观察也很难正确区分,故模型出现错误识别。且当奶牛个体快速通过视野时易产生运动模糊,导致奶牛骨架和轮廓等关键信息缺失,特征提取更为困难。

图 11 为 YOLO v3 和改进 YOLO v3 模型对视场中黑白双色的 13223 号和 16099 号奶牛的识别结果。YOLO v3 模型只识别出 13223 号奶牛,未识别出位于左下角的 16099 号奶牛。图 11a 中 16099 号

奶牛运动中背部有较大幅度的弯曲变形,因此被漏识别。此外,该视频帧图像拍摄于中午,从挤奶厅侧方有强烈的太阳光照射,使很大白色区域的奶牛背部出现了高光与阴影,产生色差,交界区域边缘出现大量未经训练的特征,导致模型漏识别。而改进 YOLO v3 模型正确识别出视场中的 2 头奶牛,如图 11b 所示,表明改进 YOLO v3 模型具有良好的抗形变能力。

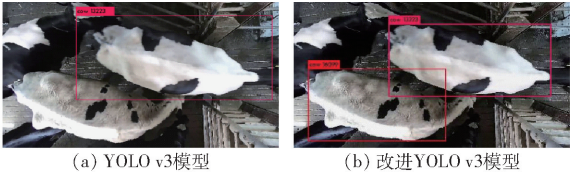


图 11 2 种模型对 13223 号和 16099 号奶牛的识别结果

Fig.11 Identification results of cow 13223 and cow 16099 with two models

未来可在不损失精度的情况下进一步优化网络模型,提高模型对奶牛的特征提取能力。也可在纯黑色、纯白色奶牛背部人为做不同标记,提高其识别特征信息的能力,进一步提高识别准确率。

4 结论

(1)在 YOLO v3 模型的基础上,借鉴 Gaussian YOLO v3 算法增加网络的输出参数,并用 K-means 算法优化奶牛图像聚类锚点框,增加残差模块以改进网络结构,提出了基于改进 YOLO v3 深度卷积神经网络的挤奶奶牛个体识别方法。

(2)在实际环境中,对 89 头奶牛的识别结果表明,改进 YOLO v3 模型对奶牛个体识别准确率为 95.91%,召回率为 95.32%,mAP 为 95.16%,平均帧率为 32 f/s,具有较高的识别准确率和适应奶牛形变的能力。

(3)本文方法识别准确率比 YOLO v3 和 Faster R-CNN 分别高 0.94 个百分点和 1.90 个百分点,检测速度是 Faster R-CNN 的 8 倍,单帧耗时仅为 0.031 s,能够满足养殖场对挤奶奶牛进行实时识别的要求。

参 考 文 献

[1] MATTHEWS S G, MILLER A L, CLAPP J, et al. Early detection of health and welfare compromises through automated detection of behavioural changes in pigs[J]. The Veterinary Journal, 2016, 217: 43-51.

[2] 何东健, 刘冬, 赵凯旋. 精准畜牧业中动物信息智能感知与行为检测研究进展[J/OL]. 农业机械学报, 2016, 47(5): 231-244.

HE Dongjian, LIU Dong, ZHAO Kaixuan. Review of perceiving animal information and behavior in precision livestock farming [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(5): 231-244. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160532&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.05.032. (in Chinese)

[3] SANTONI M M, SENSUSE D I, ARYMURTHY A M, et al. Cattle race classification using gray level co-occurrence matrix convolutional neural networks[J]. Procedia Computer Science, 2015, 59: 493-502.

[4] TSAI D M, HUANG C Y. A motion and image analysis method for automatic detection of estrus and mating behavior in cattle

- [J]. Computers and Electronics in Agriculture, 2014, 104: 25–31.
- [5] KUMAR S, PANDEY A, KONDAMUDI S, et al. Deep learning framework for recognition of cattle using muzzle point image pattern[J]. Measurement, 2018, 116: 1–17.
- [6] 何东健, 刘畅, 熊虹婷. 奶牛体温植入式传感器与实时监测系统设计与试验[J/OL]. 农业机械学报, 2018, 49(12): 195–202.
HE Dongjian, LIU Chang, XIONG Hongting. Design and experiment of implantable sensor and real-time detection system for temperature monitoring of cow[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12): 195–202. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20181225&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.12.025. (in Chinese)
- [7] 刘忠超, 何东健. 基于卷积神经网络的奶牛发情行为识别方法[J/OL]. 农业机械学报, 2019, 50(7): 186–193.
LIU Zhongchao, HE Dongjian. Recognition method of cow estrus behavior based on convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 186–193. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190719&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.019. (in Chinese)
- [8] YAJUVENDRA S, LATHWAL S S, RAJPUT N, et al. Effective and accurate discrimination of individual dairy cattle through acoustic sensing[J]. Applied Animal Behaviour Science, 2013, 146(1–4): 11–18.
- [9] 蒋国滨. 奶牛个体识别的标记方法[J]. 饲料博览, 2018(5): 86.
JIANG Guobin. Marking method for individual recognition of cow[J]. Feed Review, 2018(5): 86. (in Chinese)
- [10] 熊双辉, 孙振华. 基于RFID和图像处理的奶牛测产系统设计[J]. 计算机应用与软件, 2018, 35(7): 267–272.
XIONG Shuanghui, SUN Zhenhua. Design of milk yield monitoring system based on RFID and image processing[J]. Computer Applications and Software, 2018, 35(7): 267–272. (in Chinese)
- [11] CAPPAL M G, RUBIU N G, NIEDDU G, et al. Analysis of fieldwork activities during milk production recording in dairy ewes by means of individual ear tag (ET) alone or plus RFID based electronic identification (EID)[J]. Computers and Electronics in Agriculture, 2018, 144: 324–328.
- [12] 汪文华. 基于身体特征图像及深度学习的奶牛身份识别方法的研究[D]. 呼和浩特: 内蒙古大学, 2019.
WANG Wenhua. Research on cow identification method based on body feature image and deep learning[D]. Hohhot: Inner Mongolia University, 2019. (in Chinese)
- [13] KUMAR S, SINGH S K, SINGH R S, et al. Real-time recognition of cattle using animal biometrics[J]. Journal of Real-Time Image Processing, 2016, 13: 505–526.
- [14] 任晓惠, 刘刚, 张森, 等. 基于支持向量机分类模型的奶牛行为识别方法[J/OL]. 农业机械学报, 2019, 50(增刊): 290–296.
REN Xiaohui, LIU Gang, ZHANG Miao, et al. Dairy cattle's behavior recognition method based on support vector machine classification model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 290–296. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2019s045&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.S0.045. (in Chinese)
- [15] FUMIO O, SAYA I, YASUSHI M, et al. RGB–D video-based individual identification of dairy cows using gait and texture analyses[J]. Computers and Electronics in Agriculture, 2019, 165: 1–9.
- [16] 陈娟娟, 刘财兴, 高月芳, 等. 基于改进特征袋模型的奶牛识别算法[J]. 计算机应用, 2016, 36(8): 2346–2351.
CHEN Juanjuan, LIU Caixing, GAO Yuefang, et al. Cow recognition algorithm based on improved bag of feature model[J]. Journal of Computer Applications, 2016, 36(8): 2346–2351. (in Chinese)
- [17] ZHAO K, JIN X, JI J, et al. Individual identification of Holstein dairy cows based on detecting and matching feature points in body images[J]. Biosystems Engineering, 2019, 181: 128–139.
- [18] LI W, JI Z, WANG L, et al. Automatic individual identification of Holstein dairy cows using tailhead images[J]. Computers and Electronics in Agriculture, 2017, 142: 622–631.
- [19] CAI C, LI J. Cattle face recognition using local binary pattern descriptor[C]//Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific. IEEE, 2013: 1–4.
- [20] 张满囤, 单新媛, 于洋, 等. 基于小波变换和改进KPCA的奶牛个体识别研究[J]. 浙江农业学报, 2017, 29(12): 50–58.
ZHANG Mandun, SHAN Xinyuan, YU Yang, et al. Research of individual dairy cattle recognition based on wavelet transform and improved KPCA[J]. Acta Agriculturae Zhejiangensis, 2017, 29(12): 50–58. (in Chinese)
- [21] 赵凯旋, 何东健. 基于卷积神经网络的奶牛个体身份识别方法[J]. 农业工程学报, 2015, 31(5): 181–187.
ZHAO Kaixuan, HE Dongjian. Recognition of individual dairy cattle based on convolutional neural networks[J]. Transactions of the CSAE, 2015, 31(5): 181–187. (in Chinese)
- [22] ANDREW W, GREATWOOD C, BURGHARDT T. Visual localisation and individual identification of holstein friesland cattle via deep learning[C]//IEEE International Conference on Computer Vision Workshop Venice, Italy, 2017.
- [23] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779–788.
- [24] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2017:6517–6525.
- [25] REDMON J, FARHADI A. YOLO v3: an incremental improvement[C]//IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2018, arXiv:1804.02767.

- [26] JIANG B, WU Q, YIN X, et al. FLYOLOv3 deep learning for key parts of dairy cow body detection[J]. Computers and Electronics in Agriculture, 2019, 166: 1–8.
- [27] TOMASI C, MANDUCHI R. Bilateral filtering for gray and color images [C] // Computer Vision, Sixth International Conference on, IEEE, 1998: 839–846.
- [28] 于舒春, 臧希喆, 高永生, 等. 一种图像畸变的显-隐式校正方法[J]. 哈尔滨工业大学学报, 2008, 40(9): 1387–1390.
YU Shuchun, ZANG Xizhe, GAO Yongsheng, et al. An explicit-implicit correction method for image distortion[J]. Journal of Harbin Institute of Technology, 2008, 40(9): 1387–1390. (in Chinese)
- [29] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55(12): 20–36.
ZHENG Yuanpan, LI Guangyang, LI Ye. Survey of application of deep learning in image recognition [J]. Computer Engineering and Applications, 2019, 55(12): 20–36. (in Chinese)
- [30] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.
- [31] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137–1149.
- [32] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2016: 21–37.
- [33] CHOI J, CHUN D, KIM H, et al. Gaussian YOLOv3: an accurate and fast object detector using localization uncertainty for autonomous driving[C] // IEEE International Conference on Computer Vision, 2019: 502–511.
- [34] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.

(上接第 198 页)

- [11] HASSAAN O, NASIR A K, ROTH H, et al. Precision forestry: trees counting in urban areas using visible imagery based on an unmanned aerial vehicle[J]. IFAC-PapersOnLine, 2016, 49(16): 16–21.
- [12] HAMUDA E, GLAVIN M, JONES E. A survey of image processing techniques for plant extraction and segmentation in the field[J]. Comput. Electron. Agric., 2016, 125: 184–199.
- [13] HERNANDEZ J L, GARCÍA-MATEOS G, GONZÁLEZ-ESQUIVA J M, et al. Optimal color space selection method for plant/soil segmentation in agriculture[J]. Comput. Electron. Agric., 2016, 122: 124–132.
- [14] WOEBBECKE D, MEYER G E, VONBARGEN K, et al. Shape features for identifying young weeds using image analysis[J]. Trans. Am. Soc. Agric., 1995, 38: 271–281.
- [15] MEYER G E, HINDMAN T W, LAKSMI K. Machine vision detection parameters for plant species identification[J]. Proceedings of SPIE—The International Society for Optical Engineering, 1999: 327–335.
- [16] 翟瑞芳, 方益杭, 林承达, 等. 基于高斯 HI 颜色算法的大田油菜图像分割[J]. 农业工程学报, 2016, 32(8): 142–147.
ZHAI Ruifang, FANG Yihang, LIN Chengda, et al. Segmentation of field rapeseed plant image based on Gaussian HI color algorithm[J]. Transactions of the CSAE, 2016, 32(8): 142–147. (in Chinese)
- [17] HAMUDA E, MCGINLEY B, GLAVIN M, et al. Automatic crop detection under field conditions using the HSV colour space and morphological operations[J]. Comput. Electron. Agric., 2017, 133: 97–107.
- [18] HARALICK R M. Statistical and structural approaches to texture[J]. Proceedings of the IEEE, 1979, 67(5): 786–804.
- [19] BAKHSHIPOUR A, JAFARI A, NASSIRI S M, et al. Weed segmentation using texture features extracted from wavelet sub-images[J]. Biosyst. Eng., 2017, 157: 1–12.
- [20] WANG A, ZHANG W, WEI X. A review on weed detection using ground-based machine vision and image processing techniques[J]. Comput. Electron. Agric., 2019, 158: 226–240.
- [21] MAHALANOBIS P C. On the generalised distance in statistics[J]. Proceedings of the National Institute of Science of India, 1936, 2(1): 49–55.
- [22] 张晓伟. 光照对果树生理生化的影响及其利用[J]. 防护林科技, 2015(1): 73–76.
ZHANG Xiaowei. Effects of light on physiology and biochemistry of fruit trees and their utilization[J]. Journal of Forest Protection Science and Technology, 2015(1): 73–76. (in Chinese)
- [23] SHANNON C E. A mathematical theory of communication[J]. Bell Labs Technical Journal, 1948, 27(4): 379–423.
- [24] WONG A K C, SAHOO P K. A gray-level threshold selection method based on maximum entropy principle[J]. Systems Man & Cybernetics IEEE Transactions on, 1989, 19(4): 866–871.
- [25] HARTIG J A, WONG M A. Algorithm AS 136: a K-means clustering algorithm[J]. Journal of the Royal Statistical Society, 1979, 28(1): 100–108.