

基于YOLO-Unet组合网络的牛只个体识别方法研究

周意, 毛宽民

引用本文

周意, 毛宽民. [基于YOLO-Unet组合网络的牛只个体识别方法研究](#)[J]. 计算机科学, 2025, 52(4): 194-201.

ZHOU Yi, MAO Kuanmin. [Research on Individual Identification of Cattle Based on YOLO-Unet Combined Network](#) [J]. Computer Science, 2025, 52(4): 194-201.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多尺度融合注意力的多视角文档图像篡改检测与定位](#)

Multi-view and Multi-scale Fusion Attention Network for Document Image Forgery Localization
计算机科学, 2025, 52(4): 327-335. <https://doi.org/10.11896/jsjcx.240100142>

[大选择性核双边网络的长尾分布医学图像分类方法](#)

Long-tail Distributed Medical Image Classification Based on Large Selective Nuclear Bilateral-branch Networks
计算机科学, 2025, 52(4): 231-239. <https://doi.org/10.11896/jsjcx.240700039>

[基于虹膜纹理感知的自适应关联学习方法](#)

Adaptive Contextual Learning Method Based on Iris Texture Perception
计算机科学, 2025, 52(4): 185-193. <https://doi.org/10.11896/jsjcx.250100022>

[基于边缘约束和改进Swin Unetr的复杂器官分割方法](#)

Complex Organ Segmentation Based on Edge Constraints and Enhanced Swin Unetr
计算机科学, 2025, 52(4): 177-184. <https://doi.org/10.11896/jsjcx.240600007>

[融合词间句句多关系建模的评论推荐算法](#)

Joint Inter-word and Inter-sentence Multi-relationship Modeling for Review-based Recommendation Algorithm
计算机科学, 2025, 52(4): 119-128. <https://doi.org/10.11896/jsjcx.240700053>

基于 YOLO-Unet 组合网络的牛只个体识别方法研究

周意 毛宽民

华中科技大学机械科学与工程学院 武汉 430074

(zhou_yi1224@hust.edu.cn)

摘要 非接触式牛只个体识别方法在节约识别成本、简化识别流程和提升识别精度方面具有一定的优势,近年来在牛只个体识别领域有了充分的发展。但现有的研究中存在着识别准确率受环境和天气等外部因素影响过大、模型迁移训练困难等问题。针对上述问题,基于 YOLO-Unet 组合网络提出了包含 3 个模块的牛只个体识别模型。首先,根据 YOLOv5 模型构建图像提取模块,用以提取牛只面部图像;随后,采用 Unet 模型构建背景消除模块,用以去除牛只面部图像背景,以消除环境影响,进而提升模型泛化性能;最后,使用 MobileNetV3 构建个体分类模块,对经背景消除后的牛只面部图像进行分类。对背景消除模块进行了消融实验,实验结果表明,引入背景消除模块能极大地提升模型泛化性能。引入背景消除模块的模型在测试集上的识别准确率为 90.48%,相较于未引入背景消除模块的模型提升了 11.99%。

关键词: 牛只个体识别;深度学习;目标检测;语义分割;目标识别;泛化能力

中图分类号 TP183

Research on Individual Identification of Cattle Based on YOLO-Unet Combined Network

ZHOU Yi and MAO Kuanmin

School of Mechanical Science & Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

Abstract Non-contact method for cattle individual identification has advantages in cost reduction of identification, simplification of identification process and accurate identification, which has been fully developed in cattle individual identification in recent years. There are some problems in existing research, such as recognition accuracy affected by external factors (environment, weather, etc.) too much and difficult transfer learning. In view of the above problems, a cattle individual identification model with 3 modules based on YOLO-Unet combined network is proposed. Firstly, the image extraction module is constructed by YOLOv5 model to extract cattle facial images. Then, the background removal module is constructed by Unet model to remove the background of cattle facial images for eliminating the environmental impact, so as to improve the model's generalization ability. Finally, the individual classification module is constructed by MobileNetV3 model to classify the cattle facial images whose background removed. Ablation experiment is performed on the background elimination module. The experimental result shows that the background removal module can greatly improve the model's generalization ability. The recognition accuracy of the model with background removal module is 90.48%, which is 11.99% higher than that of the model without background removal module.

Keywords Cattle individual identification, Deep learning, Object detection, Semantic segmentation, Object recognition, Generalization Ability

1 引言

牛只个体识别方法旨在通过技术手段实现牛只身份的快速获取,有利于大型牛场对牛只进行个体管理和实时健康检测等工作,是我国畜牧业往牧场集成化、系统化与智能化方向发展过程中的重要组成部分。

现有的牛只个体识别方法主要包括接触式识别方法和非接触式识别方法。接触式识别方法包括永久识别、临时识别与电子识别。非接触式识别方法以计算机视觉技术为基础,

相比于接触式识别方法,具有更低的成本需求和更简便快捷的操作方式。因此,非接触式牛只识别方法正逐渐成为牛只个体识别领域的主要研究方向^[1-2]。牛只个体识别方法的分类如图 1 所示。

非接触式识别方法以计算机视觉技术为基础,以实现目标检测与分类为工作目标。目标检测在行人检测、车辆检测、监控安全和自动驾驶等领域取得了重大研究成果,其主要任务为定位某一类语义对象的实例。目标检测的主要步骤为:候选区域生成,候选区域特征提取,候选区域分类。目标检测

到稿日期:2024-01-17 返修日期:2024-05-20

基金项目:宁夏回族自治区重点研发计划(2020BBF02016,2021BFH02001)

This work was supported by the Key R&D Program of Ningxia Hui Autonomous Region(2022BBF02016,2021BFH02001).

通信作者:毛宽民(maokm@hust.edu.cn)

主要包括传统目标检测方法和基于深度学习的目标检测方法^[3-4]。

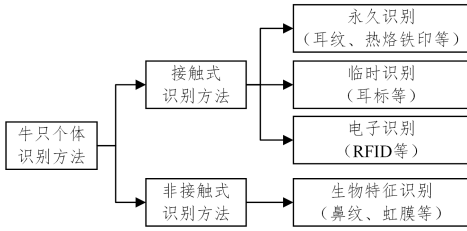


Fig.1 Classification of cattle individual identification methods

传统目标检测算法常使用不同尺寸的候选框遍历整张图像,以确定感兴趣区域的位置和范围(Region of Interest, ROI)。特征提取是指根据目标的颜色、纹理、体积等特征来设计特征提取算法,然后使用该算法对 ROI 进行特征提取。常见的特征提取算子有 Haar 算子、HOG 算子和 LBP 算子等^[5-7]。ROI 经特征提取后,需由分类器完成分类工作,如 SVM^[8]等。KUMAR 等使用 SURF 和 LBP 特征识别方法构建个体识别系统,采集牛只图像,并提取牛鼻部分图像的特征,通过度量学习与分类器集成的方法完成牛只个体识别工作^[9-10]。但牛鼻图像采集不易,相机分辨率低或拍摄位置较远时很难获取到清晰的牛鼻特征,导致识别效果较差。Chen 等提出了一种改进的特征袋模型,引入了优化后的 HOG 算子进行特征提取,自定义并构建分类器核函数,完成了牛只个体识别工作^[11]。Cai 等使用 LBP 提取纹理特征,使用多张标准牛只灰度图作为数据集对算法模型进行训练,取得了较高的准确率^[12]。

随着深度学习技术日趋成熟,非接触式牛只个体识别有了新的发展方向。Yao 等对 3 种深度网络模型进行了优化与验证,用 ResNet101 网络替代了 Faster R-CNN 算法和 R-FCN 算法中的 VGG16 网络,用 MobileNet 网络替代了 SSD 算法中的 VGG16 网络,取得了良好的检测结果。其中,优化后的 Faster R-CNN 算法的 MAP 值达到了 0.990^[13]。但此方法使用的 Faster R-CNN 算法分两步完成目标检测工作,在识别速率上有待提高。Chen 等提出了一种全局和局部网络(GPN)用于牛只个体识别工作,并进一步开发了具有空间变换的全局和局部网络(GPN-ST)模型,实验结果的 MAP 值较传统方法有 2.2% 的提升^[14]。Zhu 等提出了一种基于 CNN-SVM-ResNet 结合网络的牛脸识别模型,经训练后模型的牛只个体识别准确率达到 95.10%。但是,该模型和算法在光照强度与图像拍摄角度变化较大时,识别准确率明显下降,模型整体鲁棒性有待提升^[15]。Liu 等采用 RGB-D 图像构建数据集,使用 Inception ResNet V2 网络进行特征提取,在经过背景分离数据预处理的数据集上测试牛脸的识别准确率为 98.6%^[16]。但该方法依赖深度相机采集的深度图像,图像采集不便。

目前的非接触式牛只个体识别方法大多基于深度学习的技术实现,将牛只图像的某一部分特征作为个体分类标准,使用大量数据对识别模型进行训练。但这些方法中采集的数据大多取自相同牧场中环境相似的数个区域,具有高度相似的

背景特征;且由于背景区域占图像整体比例较大,背景特征在图像整体特征图中同样占比较大。因此,当牛只个体处于其他未在训练集数据中出现过的环境中时,识别的准确率会受到较大影响;且受背景区域影响,若需使用新的牛只个体图像数据进行个体识别工作,就需要重新训练模型,无法使用之前的模型参数作为预训练权重进行迁移训练,计算资源消耗较大。

针对以上问题,本研究提出了由图像提取、背景消去和个体识别 3 个模块组成的牛只个体识别模型。图像提取模块提取原始影像数据中的牛只面部图像,背景消去模块对数据集与待检测图像进行背景消去处理,个体分类模块完成牛只识别工作。总体工作流程如图 2 所示。

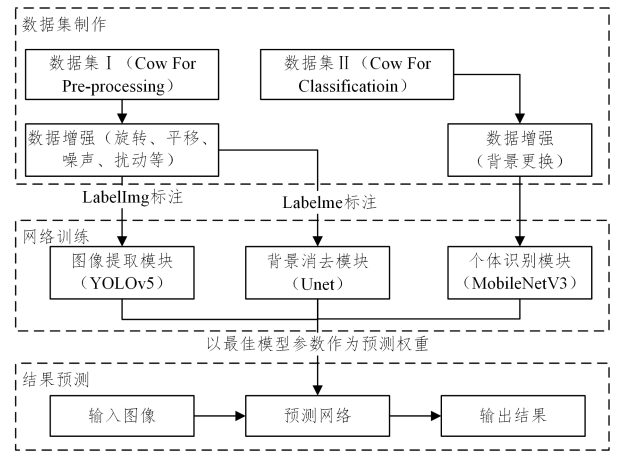


图2 牛只识别总体工作流程图

Fig.2 Overall flow chart of cattle identification

2 数据集制作

2.1 数据获取

3 个模块对数据集的要求各不相同,需单独准备与标注。图像提取模块基于 YOLOv5 模型构建,标注格式应为矩形锚框;背景消去模块的工作属于语义分割范畴,标注格式应为多边形区域;个体识别模块主要完成分类工作,需要为每张图像添加一个类别编号作为标签。

图像提取模块与背景消去模块使用开源数据集 Animals10 中的牛只面部图像部分(记为 Cow For Pre-processing, CFP)。数据集共有 1 886 张牛只图像,牛只种类、体型、所处环境及图像大小有较大差异,经特征提取网络提取出的特征比较复杂,有利于提升模型的鲁棒性。数据集中的部分图像数据如图 3 所示。



图3 部分 CFP 数据集图像

Fig.3 Partial images of CFP dataset

个体分类模块使用的数据集由 NAWFU Cowboy 团队成员前往西北农林科技大学国家肉牛改良中心实地采集(记为 Cow For Classification, CFC), 共有 11 023 张图像数据, 分别来自 20 头不同的牛只个体。数据集的部分图像如图 4 所示。



图 4 部分 CFC 数据集图像

Fig. 4 Partial images of CFC dataset

2.2 数据增强

样本数量过少和数据类型单一会对模型训练产生不利影响, 模型易出现过拟合问题。为满足各个模块使用的网络模型对数据集的需求, 需对两个数据集分别进行数据增强。

CFP 中图像类间差异较大, 前后景特征丰富, 样本复杂程度高, 图像数量较少, 需进行样本量扩充。本研究使用的数据增强策略包括: 旋转、平移、错切、翻转、裁剪、噪声注入以及亮度调整。取 CFP 中 1 686 张图像进行上述数据增强方式中的若干种处理, 得到 7 544 张图像。数据增强示例如图 5 所示。

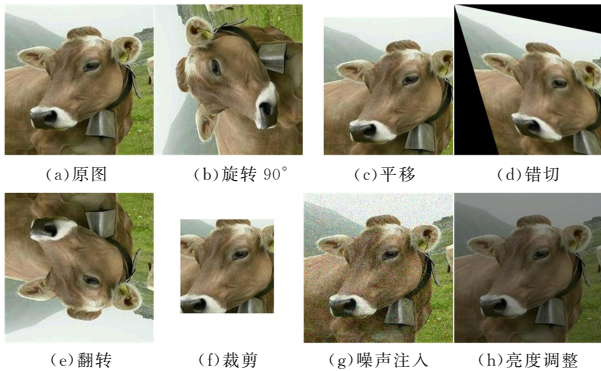


图 5 CFP 数据增强示例

Fig. 5 Data augmentation of image on CFP

CFP 样本量较大, 样本来源较少, 采集环境单一, 样本相似性较高, 需提高样本复杂程度。本研究使用背景消去模块对 CFC 中的数据进行背景消去处理, 并将不包含牛只的牧场各养殖区域图像填充至后景区域, 通过中值模糊消除前后景边缘噪点。为牛只图像配置不同的环境特征, 提升了数据集的复杂程度, 使模型注意力集中在牛只特征上, 进而提升了模型的准确率与鲁棒性。数据增强示例如图 6 所示。

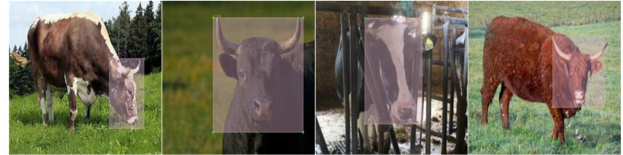


图 6 CFC 数据增强示例

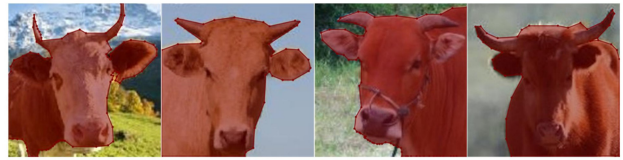
Fig. 6 Data augmentation of image on CFC

2.3 数据标注

图像提取模块基于 YOLOv5 网络构建, YOLO 数据集标注格式为 $[class, x, y, w, h]$, 其中 $class$ 表示框选区域类别, (x, y) 为框选区域中心坐标, (w, h) 为框选区域尺寸。利用 LabelImg 软件完成目标区域选取, 得到表示牛脸所在区域的标签文件。背景消去模块基于 Unet 构建, 使用语义分割通用的标注格式, 为图像中的每个像素点分配对应的分类编号。利用 Labelme 软件完成每个像素点的标注工作。标签制作结果示例如图 7 所示。



(a) LabelImg 标注示例



(b) Labelme 标注示例

图 7 标签制作结果示例

Fig. 7 Example of labels

个体分类模块旨在对输入图像进行分类, 使用的数据集为 CFC。将 CFC 中来自同一牛只的图像数据划分为一类, 建立哈希表存储图像名称与对应牛只个体名称组成的键值对, 以便快速获取图像对应牛只的分类信息。

完成标注后, 将 CFP 和 CFC 按照 8:1:1 的比例划分为训练集、验证集和测试集 3 个部分。数据集结构如 1 所列。

表 1 数据集结构设置

Table 1 Dataset structures setting

数据集名称	图像数量	归一化后尺寸	训练集数量	验证集数量	测试集数量
CFP	7 544	224 ²	6 043	755	746
CFC	11 023	224 ²	8 818	1 102	1 103

3 基于 YOLO-Unet 组合网络的牛只个体识别方法

3.1 图像提取模块

牛只养殖过程中, 躯干易沾染泥土, 从而影响个体识别准确率, 故使用面部图像作为识别标准。图像提取模块基于 YOLOv5m 网络构建, 用于从牛只图像中提取出包含面部图像的部分^[17]。YOLOv5 网络结构简图如图 8 所示。

YOLOv5 算法整体包括输入端、骨干网络 (Backbone)、颈部网络 (Neck) 和头部网络 (Head) 4 个部分。

输入端完成输入图片的预处理工作, 通过 Mosaic 数据增强方法强化模型对小目标的检测性能, 并提升模型鲁棒性; 通过对数据集进行自适应锚框计算, 获取对于当前数据集而言的最佳锚框值, 以提升模型的准确率; 通过自适应图片缩放对输入图像自适应填充最少数量的黑边, 以提升推理速度。

骨干网络用于特征提取, 一般在大型数据集 (如 COCO 和 ImageNet 等) 上完成预训练, 具备预训练参数, 输入图片在

骨干网络中经 5 次下采样后进入 SPP 模块进行多尺度特征

融合,然后进入颈部网络。

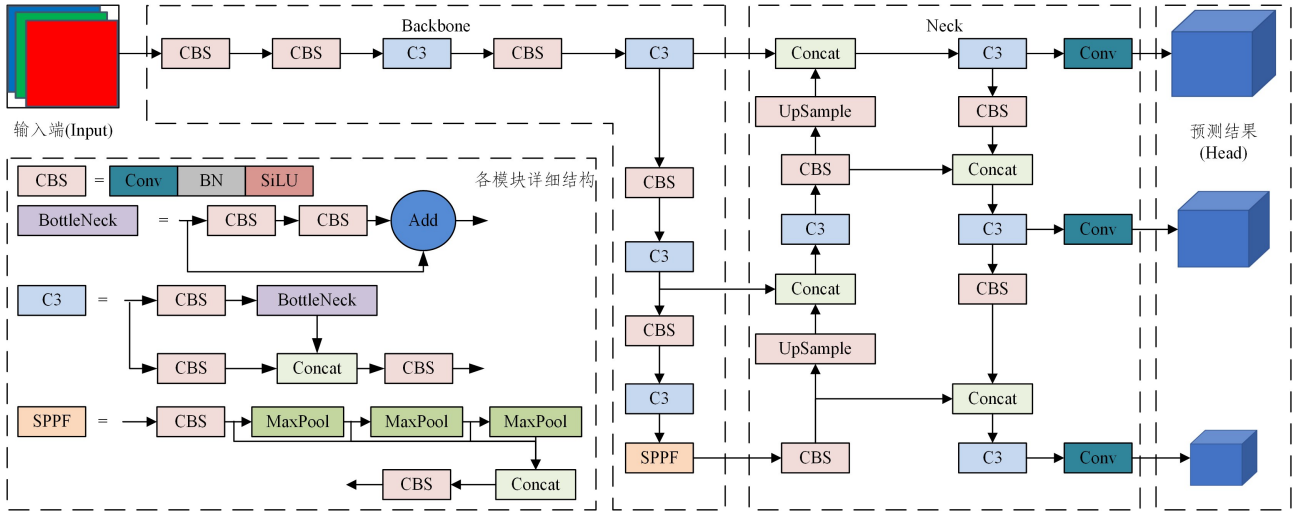


图 8 YOLOv5 网络结构简图

Fig. 8 Schematic diagram of YOLOv5 network structure

颈部网络采用 FPN+PAN 结构,其中 FPN 结构对输入进行两次上采样获取更强的语义信息,PAN 结构再进行两次下采样返回位置信息,二者拼接使得不同尺寸的特征图均包含语义信息与位置信息,保证了对不同尺寸图片的准确预测。

头部网络又称检测头,使用 CIoU_LOSS 损失函数与 NMS 非极大值抑制操作来预测目标的种类和位置。CIoU_LOSS 损失函数的计算式为:

$$L_{\text{CIoU}} = \text{IoU} - \left(\frac{\rho^2}{c^2} + \frac{v^2}{v+1-\text{IoU}} \right) \quad (1)$$

$$v = \frac{4}{\pi^2} \left(\arctan\left(\frac{w^{\text{gt}}}{h^{\text{gt}}}\right) - \arctan\left(\frac{w}{h}\right) \right)^2 \quad (2)$$

其中, L_{CIoU} 为 CIoU 损失函数; IoU 为预测框与真实框的交并比; ρ 为预测框与真实框中心坐标距离; c 为预测框与真实框角点距离的最大值; w^{gt} 和 h^{gt} 为真实框的宽与高; w 和 h 为预测框的宽与高。

3.2 背景消去模块

牛只图像经图像提取模块后,其中的牛脸部分输入到背景消去模块中进行背景消除。背景消去模块基于 Unet 网络构建,用于将牛只和背景分割成前景和背景^[18-19]。网络结构简图如图 9 所示。

Unet 网络整体结构严格对称,为编码器-解码器结构。编码器对图像进行下采样,输入图像尺寸设置为 $[512, 512]$,经 4 次下采样分别得到尺寸为 $[256, 256]$, $[128, 128]$, $[64, 64]$ 和 $[32, 32]$ 的特征图。从上至下,特征图位置信息逐渐模糊,语义信息逐渐增强。解码器对特征图进行上采样得到预测图,尺寸为 $[32, 32]$ 的特征图经一次上采样后得到尺寸为 $[64, 64]$ 的图像,将其与下采样过程中获取的同尺寸特征图进行拼接,再进行下一次上采样与拼接。重复上采样 4 次后,得到与输入图像尺寸相同的预测结果。

本研究采用交叉熵损失函数,BCEWithLogitsLoss 作为 Unet 网络的损失函数来评估预测值与真实值的相似程度,其计算式为:

$$L_{\text{BCE}} = \text{mean}\{l_0, \dots, l_{N-1}\} \quad (3)$$

$$l_n = -y_n \log(\sigma(x_n)) - (1 - y_n) \cdot \log(1 - \sigma(x_n)) \quad (4)$$

其中, L_{BCE} 为所有 batch 损失的平均值; N 为 batch 数量; n 为每个 batch 预测的图像数量; x_n 为每个维度存储一张图像的预测值; y_n 为每个维度存储一张图像的真实值; $\sigma()$ 为 Sigmoid 归一化函数。

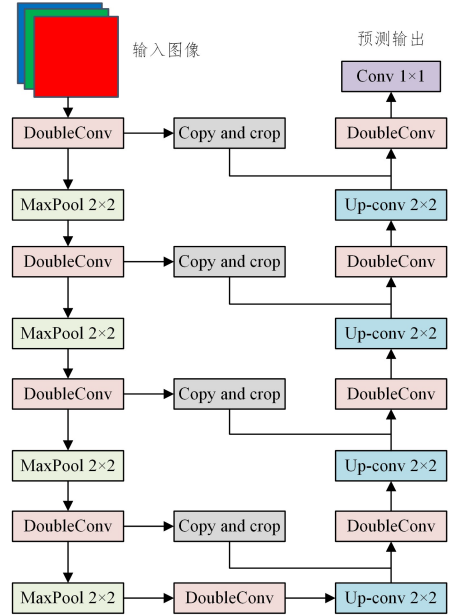


图 9 Unet 网络结构简图

Fig. 9 Schematic diagram of UNET network structure

3.3 个体分类模块

个体分类模块用于处理经背景消去后的牛只面部图像,获取对应的分类编号。为减少模型推理时间,提升模型响应速率,以轻量化的 MobileNetV3 网络为基础构建分类器^[20]。MobileNetV3 由 Google 团队提出,整体结构如图 10 所示。

MobileNetV3 网络继承了 MobileNetV1 的深度可分离卷积和 MobileNetV2 的具有线性瓶颈的残差结构,同时引入了 SE 通道注意力结构和新的激活函数 $h\text{-swish}()$,在保证准确

率的同时,通过减少计算量提升了网络计算速度。激活函数的计算式为:

$$h\text{-swish}(x) = x \cdot \frac{\text{ReLU6}(x+3)}{6} \quad (5)$$

$$\text{ReLU6}(x+3) = \min(6, \max(0, x+3)) \quad (6)$$

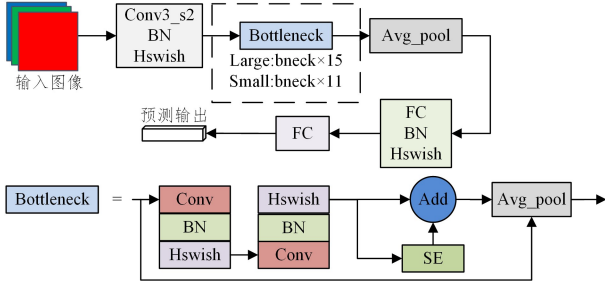


图 10 MobileNetV3 网络结构简图

Fig. 10 Schematic diagram of MobileNetV3 network structure

根据网络 bnck 模块数量、通道数以及计算量的不同, MobileNetV3 提供了两个版本的网络构建参数,分别为 MobileNetV3-large 和 MobileNetV3-small。本研究的个体分类模块采用 MobileNetV3-large 版本,牺牲了少量的检测速度,获取到了更高的检测准确性。

个体分类模块的输入为牛只面部图像,输出为一个向量,其维度等于 CFC 中牛只的种类数,向量每一维度存储的值为当前输入图像与当前维度对应的分类相同的置信度。同时,引入置信度阈值 c ,当且仅当输入图像对应的输出向量中至少有一个维度的值高于 c 时,才认为输入图像属于 CFC 中的某一类。

个体分类模块采用交叉熵损失函数 CrossEntropyLoss 作为 MobileNetV3 网络的损失函数,其计算式为:

$$\text{Loss}(\mathbf{p}, \mathbf{q}) = - \sum_x \mathbf{p}(x) \log(\mathbf{q}(x)) \quad (7)$$

其中, \mathbf{p} 为输出的预测值向量; \mathbf{q} 为 CFC 中标签的编码分布。

3.4 实验平台与模型参数

3.4.1 实验平台

本研究计算平台环境配置包括:CPU 为 Intel(R) Core (TM) i5-12400,系统运行内存为 16GB, GPU 为 NVIDIA GeForce RTX 3060,显示内存为 12GB。计算平台操作系统为 Windows 11 专业版,实验中的所有程序使用 Python 3.9 语言,基于 PyTorch 深度学习框架完成所有模块的神经网络搭建。

3.4.2 超参数设置

图像提取模块选用的超参数如下:迭代次数为 400,单次批量训练图像数量为 4,初始学习率为 0.01,模型优化器采用随机梯度下降算法 (Stochastic Gradient Descent, SGD),损失值计算采用 CIoU_LOSS 损失函数。

背景消去模块选用的超参数如下:迭代次数为 100,单次批量训练图像数量为 2,初始学习率为 1×10^{-4} ,最小学习率为 1×10^{-6} ,学习率下降采用余弦退火方式,模型优化器采用自适应矩阵估计算法 (Adaptive Moment Estimation, Adam),损失值计算采用 BCEWithLogitsLoss 损失函数。

个体分类模块选用的超参数如下:迭代次数为 500,单次批量训练图像数量为 128,初始学习率为 1×10^{-3} ,模型优化

器采用自适应矩阵估计算法 Adam,损失值计算采用 CrossEntropyLoss 损失函数。

4 实验与结果分析

4.1 评价指标

预测过程得到的所有结果都被存储到混淆矩阵 (Confusion Matrix) 中。混淆矩阵又称误差矩阵,是用于反映监督学习精度的可视化工具。多分类问题的混淆矩阵结构如表 2 所列。其中, x_{ij} 表示预测值为种类 i 的图像中,真实标签种类为 j 的图像数目, n 表示种类的总数。

表 2 混淆矩阵结构

Table 2 Confusion matrix structure

i	j				
	0	1	2	...	n
0	x_{00}	x_{01}	x_{02}	...	x_{0n}
1	x_{10}	x_{11}	x_{12}	...	x_{1n}
2	x_{20}	x_{21}	x_{22}	...	x_{2n}
...	x_{ij}	...
n	x_{n0}	x_{n1}	x_{n2}	...	x_{nm}

图像提取模块采用 P-R 曲线作为评价指标,其中 P 表示精确率, R 表示召回率,二者的计算式为:

$$\text{Precision}[i] = \frac{x_{ii}}{\sum_{j=1}^n x_{ij}} \quad (8)$$

$$\text{Recall}[i] = \frac{x_{ii}}{\sum_{i=1}^n x_{ij}} \quad (9)$$

背景消去模块采用 mIoU 值作为评价指标。mIoU 值是所有类别各自的 IoU 值求和取平均再归一化得到的结果,反映预测区域与真实区域的重合程度。

选用个体识别模块的准确率 (Accuracy) 作为整体网络性能的评估参数,其计算式为:

$$\text{Accuracy} = \frac{\sum_{i=1}^n x_{ii}}{\sum_{i=1}^n \sum_{j=1}^n x_{ij}} \quad (10)$$

4.2 实验设计

本研究完成的实验包括消融实验与对比实验两个部分。消融实验研究了模型分别在引入和不引入背景消去模块时的性能,旨在验证背景消去模块的引入有利于提升整个模型的泛化性能。针对牛只面部图像分类问题,对比实验复现了常见的用于面部图像识别的网络,旨在印证个体识别模块具有良好的分类能力。

本研究使用的数据集为自制数据集 CFP 和 CFC,分别包含 7544 和 11023 张牛只图像,以 8:1:1 的比例被划分为训练集、验证集和测试集。CFP 用于图像提取模块和背景消去模块的训练验证,个体分类模块的训练验证则在 CFC 上完成。两个自制数据集的具体制作流程参照第 2 章。

4.2.1 消融实验

本研究在数据增强操作中对 CFC 中的图像数据进行了背景更换,模拟在不同环境下采集的牛只图像,以提升数据集特征的复杂程度,进而提升模型的鲁棒性。背景消去模块旨在通过语义分割操作对分类器的训练集进行预处理,使得个体识别模块进行训练与验证时所接收的图像数据仅包含牛只

面部图像特征,提升模型对陌生图像的适应程度,进而提升模型的推理准确率与泛化性能。为验证这一猜想,本研究对其进行了消融实验。实验组 1 使用 CFP 数据集对图像提取模块进行训练验证,使用图像提取模块处理后的 CFC 数据集对个体识别模块进行训练验证。实验组 2 使用 CFP 数据集分别对图像提取模块和背景消去模块进行训练验证,使用 CFC 数据集经这两个模块依次处理后得到的图像数据作为个体识别模块的数据集,完成训练与验证。

4.2.2 对比实验

为验证个体分类模块的性能,本研究对 CNN_SVM_ResNet 组合牛脸识别网络^[15]和 ResNet50_KNN 组合牛脸识别网络^[21]进行了复现。同时,使用性能较为优秀的人脸识别模型 Facenet 进行了迁移训练^[22]。用于对比的所有模型均使用经过数据增强的 CFC 数据集,确保模型在具有良好分类能力的前提下还具有良好的鲁棒性。

4.3 结果与分析

图像提取模块的 P-R 曲线如图 11 所示。

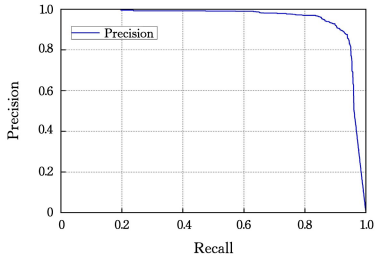


图 11 图像提取模块的 P-R 曲线

Fig. 11 P-R curve of image extraction module

二分类问题中,称 P-R 曲线与坐标轴相交得到的区域面积为 AUC 值,它反映了为每个 Recall 阈值计算所得的 Precision 的平均值。AUC 越接近 1,模型的分类性能越优秀。

对图 11 中的曲线进行积分可以得出,图像提取模块的 AUC 值为 0.95,表明该模块能较为准确地从图像中提取出牛只的面部图像部分。

背景消去模块的 mIoU 曲线如图 12 所示。

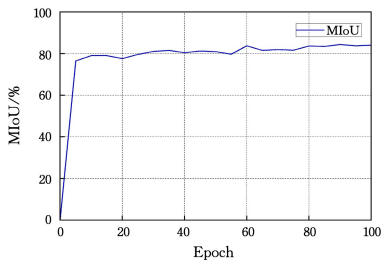


图 12 背景消去模块的 mIoU 曲线

Fig. 12 mIoU curve of background elimination module

mIoU 值越接近 1,表明预测区域与真实区域重合程度越高,mIoU 值为 1 时,表明二者完全重合。由图 12 可知,随着训练轮次的增加,mIoU 值逐渐升高,表明在语义分割的过程中,预测区域与真实区域的重合程度逐渐增长,预测错误的像素点数量逐渐减少。最后一个 epoch 的 mIoU 值为 83.43%,表明背景消去模块具有较为良好的工作性能。

两组实验的个体识别模块的准确率曲线如图 13 所示。

图中,acc_train_1 和 acc_val_1 分别为实验组 1 在训练集和数据集上的准确率曲线,acc_train_2 和 acc_val_2 分别为实验组 2 在训练集和数据集上的准确率曲线。由图 13 可知,准确率曲线较为平滑,模型准确率随训练的进行稳步提高,acc_train_1 和 acc_val_1 相较于其余二者而言,更早达到高准确率,最终准确率也略高。这说明,背景消去模块的引入小幅提升了模型的准确率,同时加快了模型训练的收敛速度。

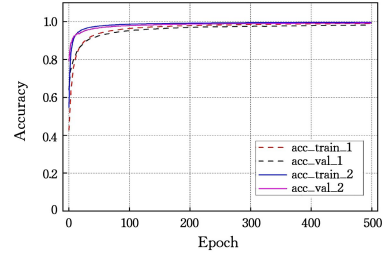


图 13 消融实验准确率曲线

Fig. 13 Accuracy curve of ablation experiment

两组实验的整体模型在训练集、验证集与全新的测试集上运行得到的结果如表 3 所列。

表 3 消融实验结果

Table 3 Results of ablation experiment

实验分组	训练集准确率	验证集准确率	测试集准确率
实验组 1	99.01	98.21	78.49
实验组 2	99.64	99.25	90.48

(%)

由表 3 实验结果可知,在测试集下,未引入背景消去模块的实验组 1 的准确率仅有 78.49%;而引入了背景消去模块的实验组 2 的准确率达到 90.48%,提升了 11.99%。这说明,背景消去模块的引入有效提升了模型的泛化性能。

待检测图像示例经 3 个模块处理后的输出如图 14 所示。输入图像经图像提取模块处理后,按得到的锚框进行裁剪,以得到牛只面部图像数据,经带自适应填充的 Resize 操作后转化为合适的大小进入背景消去模块消除背景,最后由个体识别模块完成分类工作。

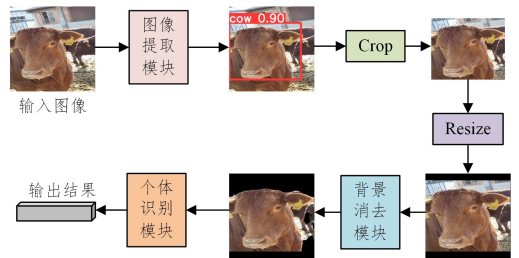


图 14 模型预测结果的可视化

Fig. 14 Visualization of model prediction results

各个模块的损失函数曲线如图 15 所示。其中,图 15(a)为图像提取模块的损失函数曲线,统计训练集的 Loss 完成于每个 epoch 的过程中,而验证集的 Loss 通常获取于每个 epoch 完成后,因此在训练前期,验证集 Loss 低于训练集为正常现象。由图 15(a)可知,损失函数曲线整体较为平滑,总体呈下降趋势,模型训练效果良好。图 15(b)为背景消去模块的损失函数曲线,曲线整体呈下降趋势,结合图 12 可知模型

具有较好的工作性能。图 15(c)和图 15(d)分别为实验组 1 和实验组 2 的个体识别模块损失函数曲线,随着训练轮次的提升,损失函数曲线整体呈下降趋势,且损失值趋近于 0,表明在 CFC 中的训练集和验证集上,两个实验组均表现出了良好的性能。

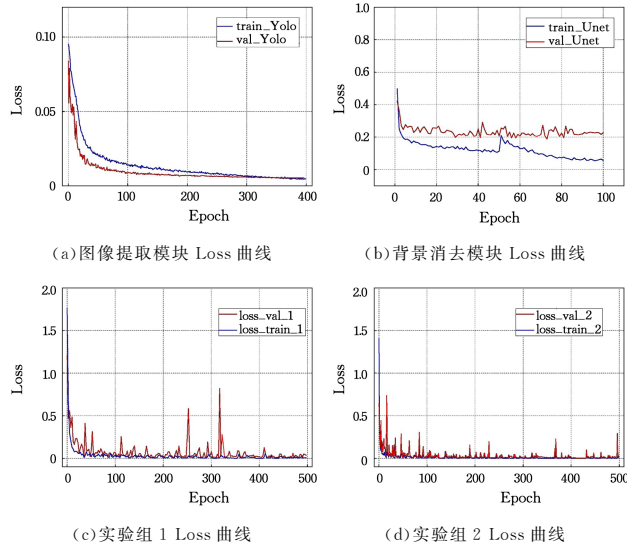


图 15 各模块的训练损失函数曲线

Fig. 15 Curves of training loss function for each module

对比实验的测试结果如表 4 所列。其中,Params 表示模型需要的参数总量,反映了模型的空间复杂度;FLOPs 表示模型完成一张图片需要进行的浮点数运算次数,反映了模型的空间复杂度;FPS 表示模型一秒内能识别的图像数量,反映了模型的整体识别速率。

表 4 对比实验测试结果

Table 4 Results of comparative experiment

分类模型	准确率/%	Params	FLOPs	FPS
CNN_SVM_ResNet	84.38	0.18×10^6	84.54×10^6	362
ResNet50_KNN	88.70	23.55×10^6	3898.32×10^6	102
InceptionResnetV1	91.94	23.49×10^6	3158.26×10^6	402
MobileNetV3	90.48	4.23×10^6	239.39×10^6	75

由表 4 可知,4 个分类模型的帧率均大于 60,能满足目前应用领域对牛只个体识别速率的需求。本研究的个体分类模块基于 MobileNetV3 网络构建,在保证较高准确率的同时,占用的空间与计算量较小,具有良好的综合性能,在移动设备软件应用开发方面具有独到的优势。

结束语 针对非接触式牛只个体识别工作中出现的模型泛化能力差、迁移训练难等问题,本研究提出了一个由图像提取、背景消去和个体分类 3 个模块组成的牛只个体识别模型。

1) 针对牛只躯干沾染泥土从而掩盖部分重要纹理特征导致的检测准确率下降问题,基于 YOLOv5 模型构建了图像提取模块,从包含牛只的图像中提取不易沾染泥土的面部部分图像,并将其输入到后续的分类器中进行特征提取与个体分类。

2) 引入了基于 Unet 语义分割网络的背景消去模块,对图像提取模块提取到的牛只面部图像进行背景消去后,再输入

到分类器中。背景消去模块在 CFP 测试集上的 mIoU 值达到了 83.43%,引入背景消去模块后,整体模型在测试集上的准确率达到 90.48%,相较于未引入时提升了 11.99%。因此,背景消去模块能在多种复杂的环境下准确提取出牛只面部图像特征,进而使模型具有良好的泛化性能。

3) 将实例分割问题拆解成 3 个步骤完成,通过增强数据集来提升特征复杂程度,并引入图像提取模块和背景消去模块着重提升模型整体的泛化性能。此外,通过引入个体识别模块着重提升模型识别的准确率。使用模型进行迁移学习时,图像提取模块和背景消去模块仅需要少量数据即可完成训练,从而减少了迁移学习的时间成本。

参考文献

- [1] XU B B, WANG W S, GUO L F, et al. A Review and Future Prospects on Cattle Recognition Based on Non-contact Identification [J]. Journal of Agricultural Science and Technology, 2020, 22(7): 79-89.
- [2] SONG Y F, WANG J, LI J L, et al. Review of individual identification methods for cattle in precision breeding mode [J]. Heilongjiang Animal Science and Veterinary Medicine, 2021, (22): 48-53, 148-149.
- [3] DU Z W, ZHOU H, LI C Y, et al. Small Object Detection Based on Deep Convolution Neural Networks; A Review [J]. Computer Science, 2022, 49(12): 205-218.
- [4] YU J Y, DING P C, WANG C. Overview: Application of Convolution Neural Network in Object Detection [J]. Computer Science, 2018, 45(S2): 17-26.
- [5] HARRIS C G, STEPHENS M. A Combined Corner and Edge Detector [C] // Alvey Vision Conference, 1988: 1-6.
- [6] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05). IEEE, 2005: 886-893.
- [7] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [8] HEARST M A, DUMAIS S T, OSUNA E, et al. Support vector machines [J]. IEEE Intelligent Systems and Their Applications, 1998, 13(4): 18-28.
- [9] KUMAR S, TIWARI S, SINGH S K. Face recognition for cattle [C] // 2015 Third International Conference on Image Information Processing (ICIIP). IEEE, 2015: 65-72.
- [10] KUMAR S, SINGH S K, DUTTA T, et al. A fast cattle recognition system using smart devices [C] // Proceedings of the 24th ACM International Conference on Multimedia, 2016: 742-743.
- [11] CHEN J J, LIU C X, GAO Y F, et al. Cow recognition algorithm based on improved bag of feature model [J]. Journal of Computer Applications, 2016, 36(8): 2346-2351.
- [12] CAI C, LI J. Cattle face recognition using local binary pattern descriptor [C] // 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. IEEE,

2013;1-4.

- [13] YAO L Y, XIONG H, ZHONG Y J, et al. Comparison of cow face detection algorithms based on deep network model[J]. Journal of Jiangsu University(Natural Science Edition), 2019, 40(2):197-202.
- [14] CHEN X L, YANG T L, MAI K Z, et al. Holstein cattle face re-identification unifying global and part feature deep network with attention mechanism[J]. Animals, 2022, 12(8):1047.
- [15] ZHU M L, ZHAO L L, HE S J. Research and realization on cattle face recognition system model based on CNN combined with SVM and ResNet[J]. Journal of Chongqing University of Technology(Natural Science), 2022, 36(7):155-161.
- [16] LIU S F, CHANG R, LI B, et al. Individual Identification of Cattle Based on RGB-D Images[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 1-10 [2024-01-12] <http://kns.cnki.net/kcms/detail/11.1964.s.20230914.1340.019.html>.
- [17] 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.
- [18] HUANG W K, TENG F, WANG Z D, et al. Image Segmentation Based on Deep Learning: A Survey[J]. Computer Science, 2024, 51(2):107-116.
- [19] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, Oc-

tober 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015:234-241.

- [20] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:1314-1324.
- [21] YANG S N, ZHAO J M, YANG M, et al. Cattle face recognition method based on debiased term SoftMax and compact metric loss function [J]. Heilongjiang Animal Science and Veterinary Medicine, 2024(4):36-42.
- [22] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015:815-823.



ZHOU Yi, born in 2000, postgraduate. His main research interests include computer vision and deep learning.



MAO Kuanmin, born in 1964, professor, Ph.D supervisor. His main research interests include machine tool dynamics and agricultural mechanization.

(责任编辑:柯颖)