doi:10.6041/j.issn.1000-1298.2024.11.010

基于特征掩膜的局部遮挡牛脸识别方法

齐咏生[1,2]　张新泽[1]　张嘉英[1,2]　刘利强[1,2]　李永亭[1,2]

(1.内蒙古工业大学电力学院,呼和浩特 010080;
2.内蒙古自治区高等学校智慧能源技术与装备工程研究中心,呼和浩特 010080)

摘要:随着智慧牧业的高速发展,牛脸识别已成为牛场智能化养殖的关键,但现实应用场景中牛脸遮挡问题较为严重,影响识别系统的性能。为此,提出一种遮挡物分割辅助牛脸识别的全新双分支网络结构。首先设计一种改进的轻量级 U-Net 遮挡物分割模型,通过加入深度可分离卷积和多尺度混合池化模块,有效提高分割网络对遮挡物的分割性能。为更好地衰减遮挡物对牛脸识别性能的影响,引入一种多级掩膜生成单元。以不同层级的遮挡分割为输入,构建识别网络不同阶段所对应的掩膜,通过掩膜运算在特征提取的各阶段有效消除遮挡造成的损坏特征信息。最后在自制数据集上进行算法有效性和实时性验证,并与多种最新的典型识别算法进行对比。实验结果表明,本文算法在遮挡牛脸数据集上平均准确率达 86.34%,识别速度为 54 f/s,且在不同程度遮挡的场景下,识别效果均优于 FaceNet 网络。

关键词:遮挡牛脸识别;图像分割;多级掩膜学习;双分支结构

中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2024)11-0093-10　　**OSID**:

Feature Mask-based Local Occlusion Cattle Face Recognition Method

QI Yongsheng[1,2]　ZHANG Xinze[1]　ZHANG Jiaying[1,2]　LIU Liqiang[1,2]　LI Yongting[1,2]

(1. *College of Electric Power*, *Inner Mongolia University of Technology*, *Hohhot* 010080, *China*
2. *Intelligent Energy Technology and Equipment Engineering Research Centre of Colleges and Universities in the Inner Mongolia Autonomous Region*, *Hohhot* 010080, *China*)

Abstract: With the rapid development of intelligent animal husbandry, bovine face recognition has become the key to intelligent cattle breeding, but the problem of bovine face occlusion in practical application scenarios is more serious, which brings challenges to the performance of the recognition system. To solve this problem, a two-branch network structure based on occluder-assisted bovine face recognition was proposed. Firstly, an improved lightweight U-Net occlusion segmentation model was designed. By adding deep separable convolution and multi-scale mixing pool module, the occlusion segmentation performance of the segmentation network was effectively improved. Secondly, in order to better attenuate the influence of occlusions on bovine face recognition performance, a multilevel mask generation unit was introduced, and masks corresponding to different stages of the recognition network were constructed with different levels of occlusions as input. The damaged feature information caused by occlusions was effectively eliminated in each stage of feature extraction through mask operation. Finally, for the validity and real-time performance of the detection algorithm, the algorithm was verified on the self-made data set, and compared with a variety of recent typical recognition algorithms. The experimental results showed that the proposed algorithm had an average accuracy of 86.34% on the blocked cow face data set, and the recognition speed was 54 f/s. Compared with the single-scale mask, the average accuracy of multistage mask was improved by 2.02 percentage points, and the recognition effect was better than that of the comparison network under different degrees of occlusion.

Key words: occluded cattle face recognition; image segmentation; multi-level mask learning; dual-branch structure

收稿日期:2024-01-31　修回日期:2024-03-07

基金项目:国家自然科学基金项目(62363029)、内蒙古科技计划项目(2021GG164)和内蒙古自然科学基金项目(2022MS06018、2021MS06018)

作者简介:齐咏生(1975—),男,教授,博士,主要从事机器人协同控制技术研究,E-mail:qys@imut.edu.cn

0 引言

近年来,随着科技的快速发展,畜牧业正朝着智能化、精细化、自动化养殖的方向发展[1]。其中,牛个体的精准识别是智能养殖的前提,牛的身份识别技术在牛个体生长记录[2]、养殖管理[3]、健康状况[4]和行为分析[5]中起着至关重要的作用,传统的牛只识别方法是在牛耳上佩戴基于射频识别技术的电子标签(Radio frequency identification, RFID)[6],但存在对牛体造成危害、人工成本高、耳标掉标和损坏、人为造假换标等问题[7]。

随着深度学习技术在视觉图像处理领域[8-10]的持续进步,基于深度学习的牛脸识别算法受到了广泛关注[11-12]。但是这些研究主要是集中在特定约束条件下的牛脸识别。

然而,在实际应用场景中,普遍存在由于牛只不受控制的行为和移动,可能导致图像遭受遮挡,进而影响识别准确率的问题[13]。为应对这一挑战,杨蜀秦等[14]提出一种融合坐标信息的奶牛面部识别模型。在 YOLO v4[15]网络的特征提取层和检测头部分分别引入坐标注意力机制和包含坐标通道的坐标卷积模块,增强模型对目标位置的敏感性,从而提升遮挡条件下牛脸识别精度。LI 等[16]设计了一个高精度的注意力机制模块(Efficient channel attention, ECA),用于提升模型区分遮挡和未遮挡部分的能力,并训练神经网络忽略被遮挡部分的特征信息。该机制被集成到 MobileNet 中,以增强遮挡条件下的牛脸识别效果。注意力机制通过给予不同区域或特征不同的权重来实现间接聚焦[17]。这种间接性导致在直观性和定位准确性方面存在一定局限,特别是在需要进行精确空间定位的遮挡识别应用场景中。

在处理遮挡识别任务研究中,除了引入注意力机制外,还常应用分割生成掩膜的方法。掩膜学习通过生成掩膜直接作用于图像中的关键遮挡区域,提供了更直观的空间信息,有助于准确定位和识别遮挡对象[18]。在遮挡人脸识别中,ZENG 等[19]提出了一种结合面部识别网络和遮挡分割网络的模型,并通过遮挡掩膜适配器模块作为桥梁,以学习不受遮挡影响的特征。QIU 等[20]提出的面部识别网络(Face recognition with occlusion masks, FROM),通过子网络动态解码准确的特征掩膜,确定输入人脸图像中因遮挡导致的损坏特征位置。并通过将提取的特征与学习的掩膜相乘,有效移除遮挡损坏区域特征。SONG 等[21]提出了一种基于成对差分孪生网络(Pairwise differential siamese network, PDSN)的掩膜学习策略。这种方法通过建立遮挡面部区域与损坏特征元素之间的对应关系,有效地提高了遮挡环境下的面部识别鲁棒性。目前未见基于掩膜学习的遮挡牛脸识别算法研究报道。

针对上述问题,本文提出一种基于特征掩膜学习策略的遮挡牛脸识别方法。模型由遮挡物分割、牛脸特征提取、掩膜学习单元 3 部分组成。不同于 FROM 和 PDSN 等单次掩膜运算来提升识别准确度,本文方法在特征提取网络的浅层、中间层、深层 3 种不同层级分别通过掩膜学习单元生成不同阶段的递进掩膜。浅层的掩膜操作帮助网络在早期阶段就过滤掉干扰信息,而在中间层和深层进一步的掩膜操作有助于细化对高级特征的理解,以有效提升在复杂实际遮挡场景中牛脸识别准确度。

1 遮挡牛脸识别算法

1.1 算法总体框架

针对遮挡牛脸识别问题,设计了一种采用遮挡物分割辅助的牛脸识别双分支结构。首先提出一种改进轻量级 U-Net 遮挡物分割网络结构。在基础 U-Net 网络中引入深度可分离卷积和多尺度混合池化模块,并采用了结合真实遮挡与模拟遮挡的混合数据集进行训练调优。其次,在识别分支的处理过程中,构建以 InceptionV3[22]为主干特征提取网络的 FaceNet 牛脸识别模型。通过 L2 正则化将提取的特征映射至 128 维的向量空间中。并采用三元组损失函数[23]进行训练调优。为更好地解决牛脸遮挡对识别性能的影响,设计了多级掩膜生成单元。将分割网络输出与识别网络相连接,使识别网络能够更有效地抑制遮挡特征。并更加关注非遮挡的牛脸部分特征。最后,在测试阶段中网络获取输入图像中的牛脸特征并与牛脸数据库中所有的牛脸进行比较,计算距离。如果距离小于门限值,则认为其具有一定的相似度。遍历数据库获得最相似的牛脸序号。判断这个序号对应的牛脸距离是否小于门限,是则认为牛脸识别成功。算法总体框架如图 1 所示。

1.2 改进轻量级 U-Net 遮挡物分割模型

U-Net 是一种用于图像分割任务的卷积神经网络架构[24]。其独特的 U 形结构包括编码器和解码器,通过跳跃连接将不同层级的特征图相互关联,以有效学习局部和全局特征。最终生成与输入图像相同大小的分割掩模。本文中用深度可分离卷积代替标准卷积,减少模型计算复杂度和参数量。并在编码器和解码器跳跃连接部分添加

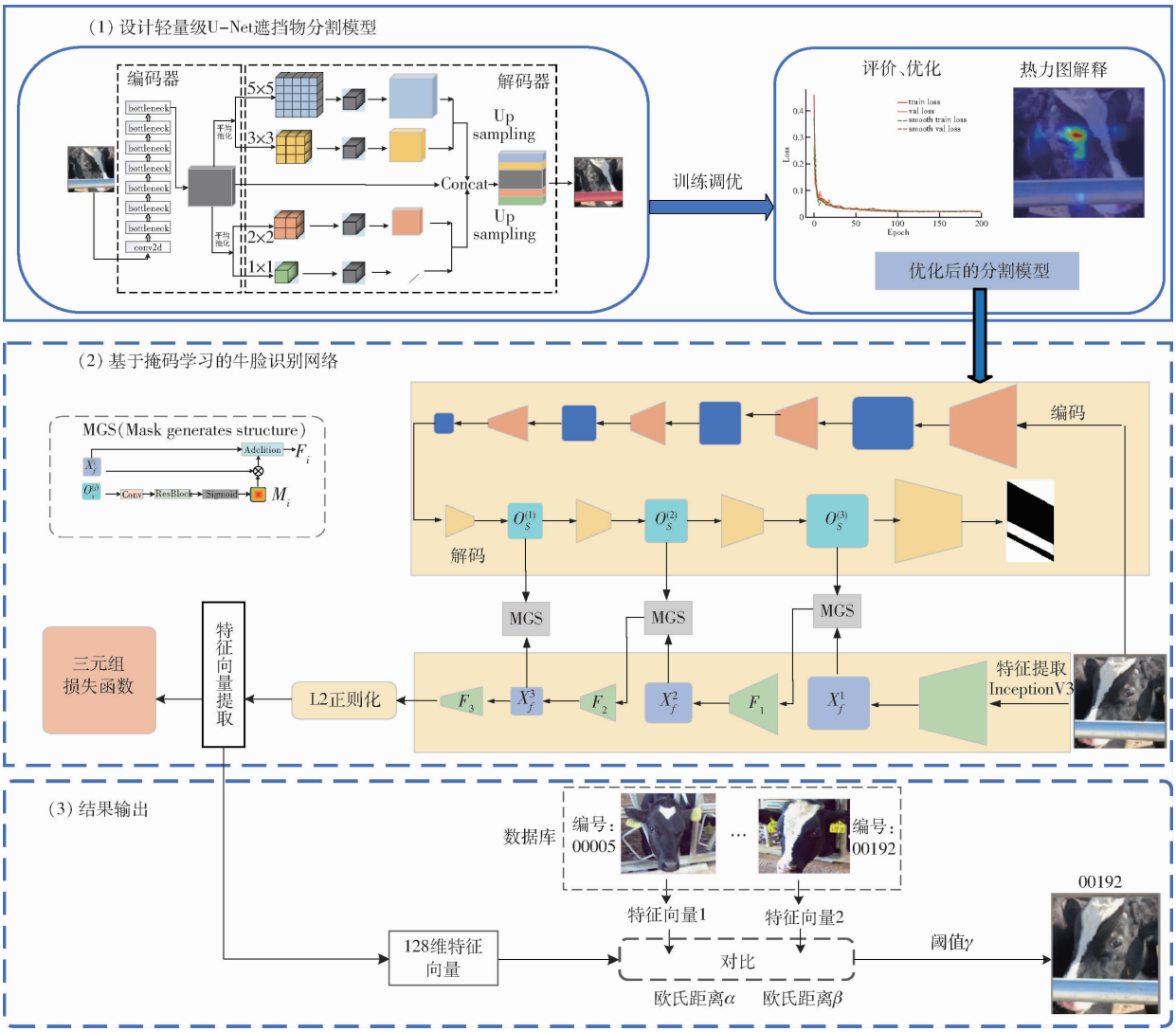

图 1　算法结构框图

Fig. 1　Algorithm framework

多尺度混合池化模块，以进一步提升图像分割的性能。

1.2.1　深度可分离卷积

深度可分离卷积（Depthwise separable convolution，DSC）是目前轻量化网络中应用最广泛的一种卷积方式。其主要作用是减少网络参数，从而提高网络推理速度。相对于传统卷积（图 2a），可分离卷积用于提高算法运行速度主要分为两步：逐通道卷积和逐点卷积，示意图如图 2b、2c 所示。

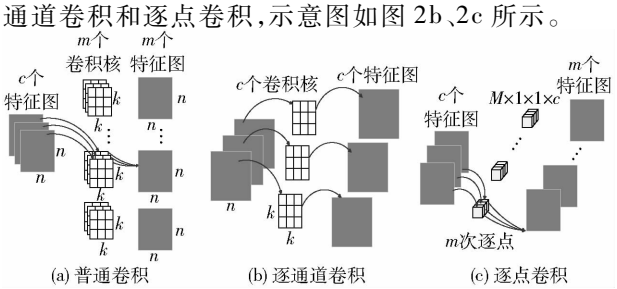

图 2　通道卷积示意图

Fig. 2　Channel convolution schematic

假设输入特征图尺寸为 $n \times n \times c$，卷积核尺寸为 $k \times k \times m$，输出特征图尺寸为 $n \times n \times m$，标准卷积层参数量为

$$C_{\text{normal}} = cmk^2 \tag{1}$$

深度可分离卷积一般分为两个步骤：首先对输入图像或特征图进行逐通道卷积，然后对逐通道卷积得到的特征图进行逐点卷积。对于相同输入，DSC 参数量计算式为

$$C_{\text{DSC}} = ck^2 + cm \tag{2}$$

其中，第 1 部分为逐通道卷积参数量 ck^2，第 2 部分为逐点卷积参数量 cm。所以深度可分离卷积与标准卷积参数量之比为

$$\frac{C_{\text{DSC}}}{C_{\text{normal}}} = \frac{ck^2 + cm}{cmk^2} = \frac{1}{m} + \frac{1}{k^2} \tag{3}$$

显然 $C_{\text{DSC}} < C_{\text{normal}}$，即深度可分离卷积的卷积核数量少，运行速度更快。

1.2.2　多尺度混合池化模块

在对分割网络测试调优的过程中发现，U－Net

网络对于结构简单的遮挡物分割效果好。对于结构更加复杂的遮挡物,如具有复杂空间分布和多样性纹理的牛颈枷,其效果有所下降。因为在较浅层,感受野相对较小,只能涵盖颈枷的局部信息。如果仅依赖于较浅层的特征,可能无法充分捕捉到遮挡物的全貌和细节。考虑到遮挡物可能存在复杂的空间结构,涉及多个尺度信息,本文提出了一种多尺度混合池化模块(Multi-scale mixed pooling module, SMPM),如图 3 所示。

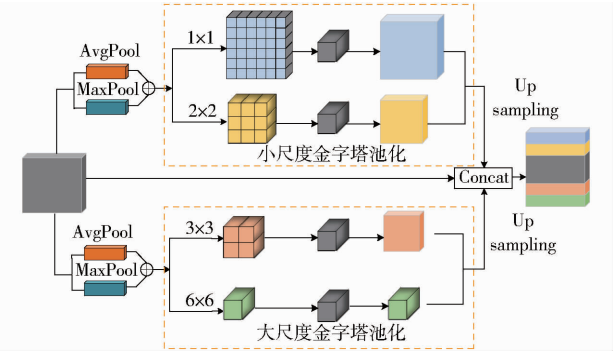

图 3 多尺度混合池化模块

Fig. 3 Multi-scale mixed pooling module

在 U-Net 的编码器和解码器的跳跃连接结构中,为了更好地捕捉多尺度信息,提高网络对遮挡物的理解和分割性能,引入多尺度混合池化模块。其中最大池化有助于保留图像中最显著的特征,确保网络关注到物体边缘、纹理等重要信息。平均池化通过计算池化窗口内的平均值,有助于平滑图像,减轻噪声,提取更整体的特征信息。混合池化模块是将输入特征图以并行的方式进行平均和最大池化[25],在通道维度上进行拼接。多尺度池化模块通过采用 4 种不同尺寸的并行池化操作,即 1×1、2×2、3×3 和 6×6,获取多个尺寸的特征图。其中,1×1 和 2×2 池化操作用于提取大尺度物体的特征,而 3×3 和 6×6 池化操作用于捕捉小尺度物体的特征。接着,通过应用"1×1 卷积"来降低通道数。最后,使用双线性插值进行上采样,以获得与金字塔模块之前相同尺寸的特征图,并将它们在通道维度上拼接在一起。

1.3 多级掩膜生成单元

关于人脸遮挡识别领域研究的文献[20]证明:使用二值化的掩膜会完全去除遮挡区域的特征,但这些特征仍包含部分用于识别的次要信息。为了测试完全去除遮挡区域对牛脸识别的影响,首先用改进 U-Net 网络将遮挡牛脸图像生成二值化掩膜,其中遮挡物部分值为 0,牛脸部分值为 1。通过逐元素相乘,得到去除遮挡的牛脸图像。使用识别网络将牛脸图像提取为 128 维特征向量。其中每个维度对应于某种牛脸特征。对于两个特征向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,它们之间的欧氏距离计算式为

$$Distance(\boldsymbol{A},\boldsymbol{B}) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2} \tag{4}$$

式中 n——牛脸图像特征向量的维度

A_i——特征向量 $\boldsymbol{A}$ 第 i 个维度

B_i——特征向量 $\boldsymbol{B}$ 第 i 个维度

欧氏距离越小表示它们在特征空间中越相似,越可能属于同一头牛,值越大越可能是不同的牛。如图 4 所示,相同牛在完全去除遮挡物的情况下,欧氏距离从 0.755 增大到 0.932,不同牛之间测试在完全去除遮挡物的情况下与未经处理的原图相比较,欧氏距离从 1.513 下降到 1.392。实验结果显示直接使用二值化的掩膜完全去除遮挡区域,反而会降低牛脸识别正确率。因为这些遮挡区域的特征仍包含部分用于识别的边缘轮廓等次要信息。

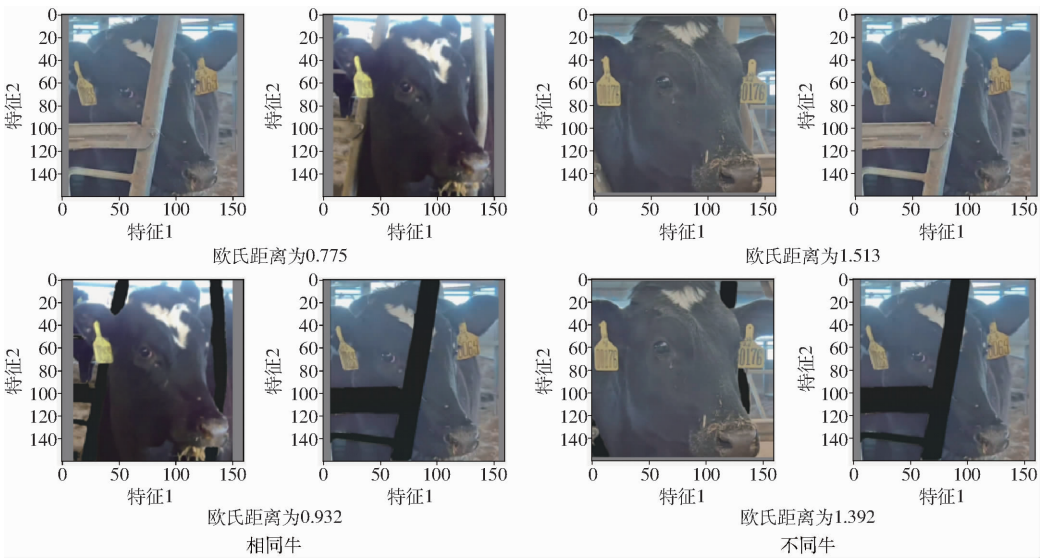

图 4 完全移除遮挡的牛脸识别结果

Fig. 4 Recognition of cattle faces with complete obstruction removal

因此，本文提出多级掩膜生成单元，针对不同层次的特征，生成不同尺度的潜在掩膜。对遮挡特征采取抑制的策略而非完全去除的方法以消除其影响。掩膜学习单元（Mask generates structure，MGS）结构如图 5 所示。

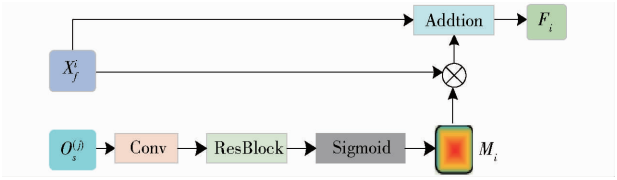

图 5　MGS 结构图

Fig. 5　MGS structure diagram

改进 U－Net 网络的解码器生成层次化的遮挡分割 $O_S=\{O_s^{(1)},O_s^{(2)},O_s^{(3)}\}$。将 $O_s^{(j)}$ 输入到掩膜生成单元中，经过掩膜学习得到非二值化的多级掩膜 $M_i(i=1,2,3)$，掩膜学习过程为

$$M_i=\text{Sigmoid}(\theta(\text{Conv}(O_s^{(j)}))) \tag{5}$$

式中　Conv(·)——卷积操作

$\theta(\cdot)$——残差块运算符

对输入进行残差块的运算以增强 MGS 的学习能力。通过 Sigmiod 运算将特征掩膜中的元素值约束在(0,1)。特征掩膜 M_i 与原始特征 X_f^i 相乘即可掩蔽掉被遮挡损坏的特征，MGS 模块最后输出的 F_i 计算过程为

$$F_i=\begin{cases}X_f^i\otimes M_i & (i=1)\\ X_f^i+X_f^i\otimes M_i & (i=2,3)\end{cases} \tag{6}$$

式中　X_f^i——第 i 级特征提取块生成的特征

采用的多级掩膜运算，多个掩膜生成单元共享相同的网络架构，后两级输入 X_f^i 同样是经过前一层的掩膜运算后得到，F_i 因此在生成中引入残差连接，将原始输入 X_f^i 与中间层特征 $X_f^i\otimes M_i$ 相加。若输入牛脸未被遮挡，那么掩膜生成器输出的特征掩膜中的元素值将趋近于 1，即 $M_i\approx 1$，从而使模型保持对无遮挡牛脸的通用性。

2　实验

为满足后期算法可移植性的要求，算法首先在服务器上进行训练。训练环境均在 Ubuntu 18.04 Linux 操作系统下完成，硬件环境配置为：i7－11700 @2.50 GHz CPU、NVIDIA GeForce RTX 3090 显卡、24 GB 运行内存。软件环境配置为：Cuda 11.0、Cudnn 8.1.0 加速，基于 pytorch 1.8.1 深度学习框架。

模型训练策略：采用两阶段训练方式，先采用交叉熵损失函数训练分割网络为

$$L_{occ}=-\sum\left(y\log(p)+(1-y)\log(1-p)\right) \tag{7}$$

式中　y——真实标签

p——模型预测为正类的概率

使用 Adam 优化器进行优化，学习率设为 0.001，迭代动量系数为 0.9，批量大小为 32，迭代次数为 300。

在训练识别网络时，将训练好的分割网络输出用作识别网络训练的辅助信息，并确保分割网络的权重被冻结，以防止之前训练好的分割模型在后续训练中性能退化。识别网络采用三元组损失函数进行训练，计算式为

$$L_{cls}=\max(\text{dist}(A,P)-\text{dist}(A,N)+Margin,0) \tag{8}$$

式中　dist(A,P)——Anchor 和 Positive 之间距离

dist(A,N)——Anchor 和 Negative 之间距离

Margin——超参数为 Positive 和 Negative 之间最小距离

使用 Adam 优化器进行优化，迭代动量系数为 0.9，批量大小为 32，总迭代次数为 500。

2.1　数据采集

采用海康威视 F－LZS 型智能摄像机，分辨率为 1 440 像素 ×1 080 像素，帧率为 30 f/s。以该型号摄像机为基础，设计单机多点架设的方式对呼和浩特市林格尔县欧太牧业的 230 头荷斯坦奶牛、38 头安格斯肉牛、12 头黄牛进行跟踪拍摄，共有单机设备 6 套，分别布设通过 8 路 POE 录像机整体控制，用于采集分割数据和识别数据。

2.2　数据集制作

在牛脸视频数据处理中，采用每 0.8 s 提取一帧的方法初步筛选关键帧，以保证图像多样性。针对牛长时间静止或微小运动造成的重复图像，进一步应用基于结构相似性指数（SSIM）的去重技术，有效移除相差微小的重复牛脸图像。利用 Faster R－CNN 将图像中牛脸图像剪裁并保存。得到 13 800 幅遮挡牛脸图像以及 2 800 幅无遮挡牛脸图像。

分割数据集：选取 3 300 幅包括不同遮挡程度和不同遮挡位置的牛脸图像，并在无遮挡牛脸图像上添加了树叶、枝干等遮挡物，制作了含 900 幅合成遮挡牛脸图像的数据集来增加遮挡物的多样性，避免分割网络因遮挡种类单一而产生过拟合。使用交互式自动标注工具 X－AnyLabeling 生成遮挡物的标注掩膜[26]，得到 json 文件并将其转换为 8 位灰度图作为标签，如图 6 所示，再将图像数据集按比例 8∶1∶1 随机划分为训练集、验证集、测试集。

识别数据集：选取 280 头牛多种角度、不同光照情况下的牛脸图像，其中包含 12 700 幅遮挡牛脸图

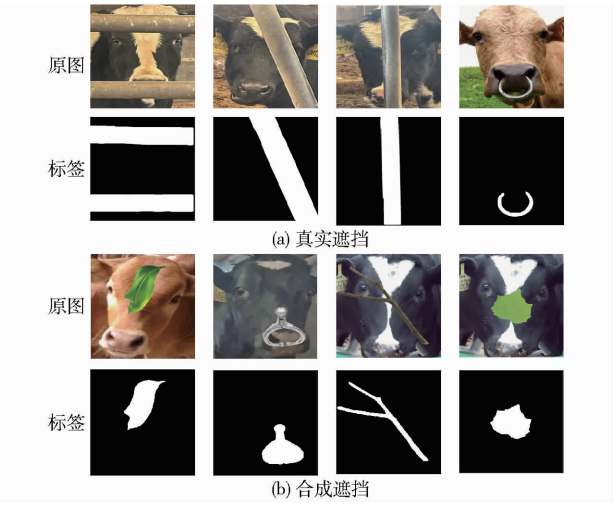

图 6 分割数据集标签

Fig. 6 Segmentation dataset labels

像和 2 400 幅无遮挡牛脸图像。无遮挡的牛脸图像可以帮助识别网络学习牛脸的基本特征,如眼睛、耳朵、鼻子位置和形状。为算法提供了一个基准,以便更好地理解和识别在遮挡条件下的牛脸。为了保证最终能够有效地提供给网络以供训练,将牛脸图像进行不失真的图像放缩,转换为 284 像素 × 284 像素的图像。数据具体划分数量为:分割训练集 3 600 幅、分割验证集 300 幅、分割测试集 300 幅、识别训练集 12 700 幅、识别验证集 1 600 幅、识别测试集 800 幅。

2.3 性能评价标准

实验使用平均交并比(Mean intersection over union, MIoU)和平均像素准确率(Mean pixel accuracy, MPA)作为衡量遮挡物分割模型性能评价指标,使用识别准确率(Accuracy, Acc)作为遮挡牛脸识别模型性能评价指标。Acc 是模型正确判断同一牛脸样本与正确判断不同牛脸样本在总样本中所占的比例。并通过模型参数量和帧率(Frame per second, FPS)对识别模型在实际应用中的可行性进行评估验证。

2.4 算法测试

2.4.1 分割算法消融实验

在分割数据集上进行消融实验,以评估每项改进的效果。实验结果如表 1 所示。引入深度可分离卷积(DSC)后,模型内存占用量显著减少,从 37.98 MB 降至 25.51 MB,并且帧率有所提高,这证明 DSC 在减少模型参数和提高推理速度方面具有显著效果。但分割准确率却下降约 0.5 个百分点。当引入多尺度最大池化时,模型平均交并比(MIoU)和精确度均得到了提升,反映了最大池化在保留图像边缘和纹理特征方面的优势。而采用多尺度平均池化同样改善了模型性能,这表明平均池化有助于平滑图像和降低噪声,有利于全局特征的提取。最终,融合多尺度混合池化的模型在 MIoU 和精确度上均展现出最佳表现,与仅采用最大池化或平均池化的模型相比,交并比分别提高 0.62 个百分点和 0.49 个百分点。尽管在速度方面略慢于单独使用 DSC 的模型,但其处理速度仍优于 U-Net 模型。

表 1 改进 U-Net 分割网络消融实验结果

Tab. 1 Improved U-Net segmentation network ablation experiment

U-Net	深度可分离卷积	多尺度平均池化	多尺度最大池化	多尺度混合池化	平均交并比/%	平均像素精准率/%	内存占用量/MB	帧率/(f·s^{-1})
√					87.24	90.13	37.98	48
√	√				86.70	89.63	25.51	64
√	√	√			88.56	91.21	31.24	62
√	√		√		88.69	91.33	31.33	62
√	√			√	89.18	91.72	32.51	61

注:"√"表示选用该组件。

为验证本文提出的分割算法在不同遮挡情况下的鲁棒性,选取简易遮挡、复杂遮挡和合成遮挡 3 种不同的样本进行测试。图 7 展示了不同遮挡条件下对比测试结果。从图中可以观察到,对于简单遮挡,分割网络都能有效地识别并分割出遮挡物。在处理形状结构复杂的遮挡物分割任务时,融合了深度可分离卷积和多尺度混合池化的分割模型在整体和细节的分割精度上均超过了其他模型。该模型展现了较低的误识率,并能够精细地提取遮挡物边缘的细小特征,同时保持了分割边缘的光滑性和清晰度。

2.4.2 特征提取网络对比

为选择最佳的主干特征提取网络,分析对比 InceptionV3、Inception ResNetV1[27]、ResNet50[28]、MobileNetV2[29]4 种被广泛应用在面部识别任务中的特征提取网络,并且在相同训练参数设置下,对无遮挡和有遮挡牛脸测试集进行准确率实验,结果如表 2 所示。

由表 2 可以看出,基于 InceptionV3 主干网络对无遮挡和有遮挡测试集识别准确率分别为 96.01% 与86.34%,与参数量相似的 Inception

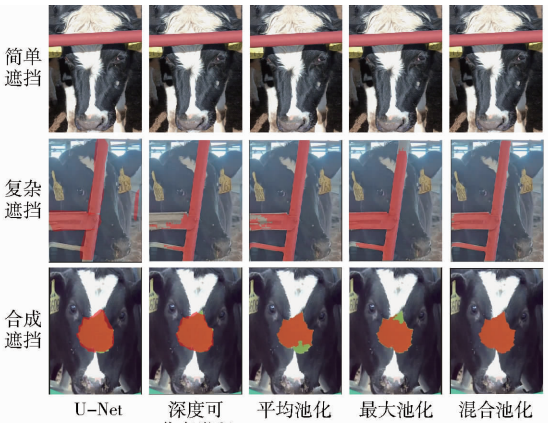

图 7　多种遮挡条件下分割效果对比

Fig. 7　Comparative segmentation under various occlusion conditions

表 2　不同特征提取网络对比实验结果

Tab. 2　Comparative experiment of different feature extraction networks

网络	无遮挡准确率/%	有遮挡准确率/%	参数量
InceptionV3	96.01	86.34	2.38×10^7
Inception ResNetV1	95.27	85.75	2.75×10^7
ResNet50	95.43	86.01	2.51×10^7
MobileNetV2	91.14	80.82	3.42×10^5

ResNetV1 和 ResNet50 网络相比，无遮挡牛脸准确率分别高 0.74、0.58 个百分点，有遮挡牛脸准确率分别高 0.59、0.33 个百分点。虽然 MobileNetV2 具有参数量少、识别速度快的优势，但是准确率偏低。实验结果表明 InceptionV3 更适合作为主干特征提取网络开展遮挡牛脸识别研究。

2.4.3　基于掩膜学习的牛脸识别算法测试结果

首先定义 4 种遮挡程度以评估牛脸识别的准确性，如图 8 所示。其中轻度遮挡为遮挡区域占牛脸总面积比例小于 20%，遮挡并不完全覆盖牛脸的任何主要特征，如眼睛、鼻子或嘴巴。中度遮挡指遮挡区域占比为 20% ~40%。一些主要特征可能被部分遮挡，但仍有足够区域可供识别。重度遮挡则是指遮挡区域在牛脸总面积的40% ~60%之间，大部分主要特征被遮挡，识别变得极其困难。



图 8　不同遮挡程度的牛脸图像

Fig. 8　Images of cattle faces with varying degrees of occlusion

为了验证不同遮挡程度下的牛脸网络识别准确率，在训练网络时对验证数据集进行整理，依次选取 400 幅不同遮挡程度的牛脸图像作为验证数据集，并确保未在训练数据集中出现，不同遮挡下验证集准确率收敛曲线如图 9 所示。

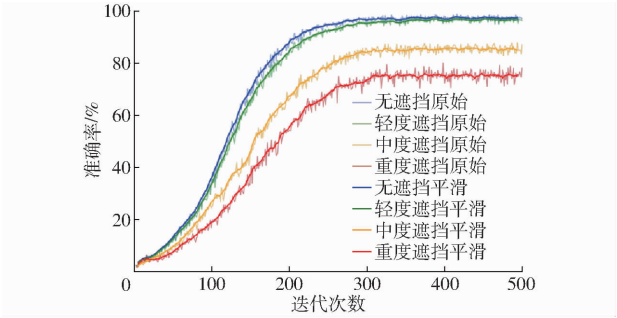

图 9　不同遮挡程度下验证集的准确率收敛曲线

Fig. 9　Convergence curves of validation set accuracy at different occlusion levels

无遮挡和轻度遮挡（蓝线和绿线）验证集在网络训练 500 次后准确率基本稳定在 96% 左右。如果输入图像未被遮挡，那么掩膜生成器输出的特征掩膜中的元素值将趋近于 1，对识别网络中生成的牛脸特征不产生影响，从而使模型保持对无遮挡牛脸识别的通用性。中度遮挡虽然一些主要特征可能被部分遮挡，牛脸验证集准确率仍然达 86%。在遮挡区域占牛脸总面积的 40% ~60% 之间的重度遮挡（红线）情况下，牛的面部特征大部分缺失，仅依靠剩余的面部信息难以达到高识别精度。验证集的准确率为 76%。

遮挡位置的不同会对识别网络带来不同的性能下降，如图 10 所示。通过对测试数据集中遮挡左

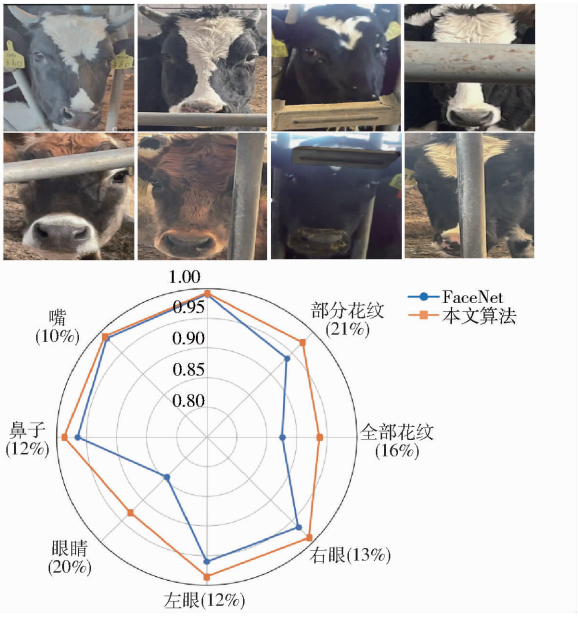

图 10　不同遮挡位置下网络识别准确率雷达图

Fig. 10　Radar chart of network recognition accuracy under different occlusion positions

眼、右眼、嘴巴、鼻子、双眼、部分花纹、全部花纹的牛脸进行归类，构建 7 个不同遮挡的牛脸数据集，并通过制作合成遮挡方式确保每组图像都有 60 幅。选用以 InceptionV3 为提取网络的 FaceNet 为基线模型与本文算法进行对比。

从图 10 可以看出，嘴巴和鼻子的遮挡对准确率的影响较小，表明在识别过程中的作用可以忽略。但是花纹遮挡和眼睛遮挡的情况下，FaceNet 的准确率仅为 87.13% 和 84.78%。本文算法识别准确率也分别下降到 94.07% 和 92.36%。表明对于牛脸识别任务，眼睛和花纹是关键的身份信息。从识别准确率雷达图可知，本文方法在所有的遮挡条件下都显著地超过 FaceNet 网络。特别是，当遮挡变得更加困难时，准确率的差距变得更大。

为进一步评估模型的优劣，将本文算法的掩膜单元进行多级掩膜与单一掩膜的对比实验，并与其他当前最新的牛脸遮挡识别算法进行对比，包括：FaceNet、TB - CNN[11]、ECANet[16]。对比测试结果如表 3 所示。

表 3 不同牛脸识别算法对比

Tab. 3 Comparison of different cattle face recognition algorithms

模型	准确率/%			平均准确率/%	帧率/(f·s^{-1})
	轻度遮挡	中度遮挡	重度遮挡		
FaceNet	92.25	78.51	63.12	77.96	58
ECANet[16]	93.55	83.78	74.36	83.89	38
TB - CNN[11]	93.78	85.12	73.98	84.29	42
本文算法(单级掩膜)	93.84	85.01	74.13	84.32	56
本文算法(多级掩膜)	96.01	86.43	76.59	86.34	54

本文算法在识别精度上显著高于其他识别算法，此外，相比于单一尺度的掩膜，采用多级掩膜平均准确率提高 2.02 个百分点。在识别算法速度方面虽然比 FaceNet 网络慢，但从图 10、表 3 可知 FaceNet 对遮挡牛脸识别准确率最低。从对比结果可以看出，本文算法在识别准确率上具有明显优势，能更好适应不同遮挡程度的牛脸识别任务。

2.4.4 特征图与特征分布可视化

为更好地说明本文多级掩膜学习方法消除遮挡区域特征影响的过程，以及遮挡检测分割模块对输入识别牛脸的处理效果，本文对改进 U - Net 提取的特征 X_i、掩膜单元生成的掩膜 M_i、掩膜运算后的特征 F_i 进行热力图可视化展示。

如图 11 所示，生成的掩膜不同于遮挡分割模块生成的二值化分割表示，并不是直接使用遮挡分割

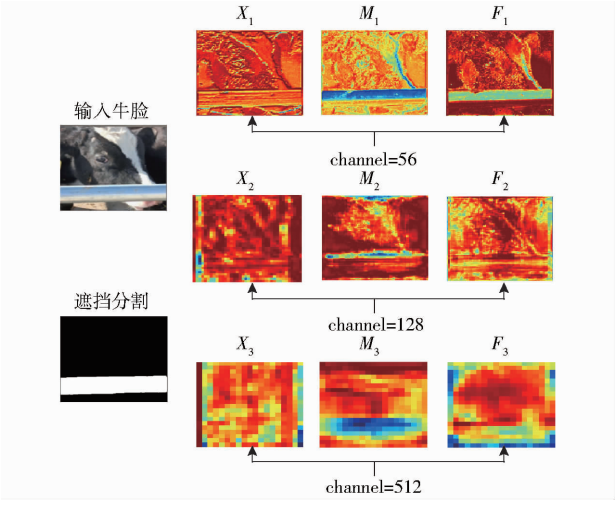

图 11 中间层特征热力图可视化

Fig. 11 Visualization of mid-layer feature heatmaps

表示图与牛脸特征相乘去除遮挡区域的全部特征，而是通过掩膜运算抑制遮挡区域的特征，这使得可用于识别的次要信息得以保留，未被遮挡的牛脸信息点得到更多的关注。

为了展示本文方法的身份分类能力，从测试集中随机选取 8 头牛，每头牛拥有 15 幅在不同遮挡条件下拍摄的面部图像。采用 t-SNE 技术对牛脸识别算法提取的 128 维特征进行降维，以便在二维空间中进行可视化分析(图 12)。生成的图像中每个点代表了特定遮挡条件下提取的嵌入式特征向量。相同颜色的点表示同一头牛的不同图像特征。与 FaceNet 算法相比，本文提出的算法在特征嵌入中展示了更小的类内距离和更大的类间距离。这表明该方法在类内一致性方面更为强大，同时也更加有效地区分不同类别。特别地，这种特征分布的变化证明了多级掩膜学习方法在提高模型对遮挡条件下的鲁棒性方面的有效性，使得算法能够更加清晰地区分不同遮挡条件下牛面部的特征。

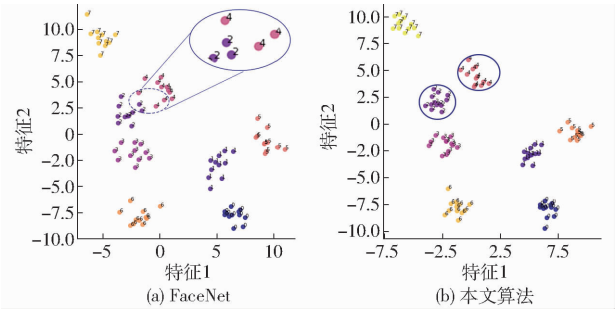

图 12 特征分布的 t-SNE 可视化

Fig. 12 The t-SNE visualization of feature distribution

3 结论

(1)将掩膜学习应用到遮挡牛脸识别任务中。

引入改进的 U－Net 遮挡物分割模型辅助牛脸特征识别的双分支网络结构，通过多级掩膜生成单元，识别网络能更有效地处理遮挡特征，专注于非遮挡的牛脸部分特征。

(2)对不同遮挡程度牛脸数据集，进行识别准确度实验。本文算法平均准确率达 86.34%，识别速度为 54 f/s。且采用多级掩膜比单一尺度的掩膜平均准确率高 2.02 个百分点。与多种最新的典型识别算法对比分析，在牛脸轻度、中度以及重度遮挡的场景中，识别效果均优于对比网络。

参 考 文 献

[1] MAHMUD M S, ZAHID A, DAS A K, et al. A systematic literature review on deep learning applications for precision cattle farming[J]. Computers and Electronics in Agriculture, 2021, 187: 1－16.

[2] 杨梅, 赵建敏. 基于三元损失的保险业务牛脸识别系统的设计[J]. 光电子·激光, 2022, 33(8): 831－839.
YANG Mei, ZHAO Jianmin. Design of cow face recognition system for insurance business based on three-dimensional loss algorithm[J]. Journal of Optoelectronics·Laser, 2022,33(8):831－839. (in Chinese)

[3] MANCUSO D, CASTAGNOLO G, PORTO S M C. Cow behavioural activities in extensive farms: challenges of adopting automatic monitoring systems[J]. Sensors, 2023, 23(8): 3828.

[4] 黄小平, 冯涛, 郭阳阳, 等. 基于改进 YOLO v5s 的轻量级奶牛体况评分方法[J]. 农业机械学报, 2023, 54(6): 287－296.
HUANG Xiaoping, FENG Tao, GUO Yangyang, et al. Lightweight dairy cow body condition scoring method based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 287－296. (in Chinese)

[5] 王政, 许兴时, 华志新, 等. 融合 YOLO v5n 与通道剪枝算法的轻量化奶牛发情行为识别[J]. 农业工程学报, 2022, 38(23): 130－140.
WANG Zheng, XU Xingshi, HUA Zhixin, et al. Lightweight recognition for the oestrus behavior of dairy cows combining YOLO v5n and channel pruning[J]. Transactions of the CSAE, 2022, 38(23): 130－140. (in Chinese)

[6] WILLIAMS L R, FOX D R, BISHOP-HURLEY G J, et al. Use of radio frequency identification (RFID) technology to record grazing beef cattle water point use[J]. Computers and Electronics in Agriculture, 2019, 156: 193－202.

[7] KUMAR S, SINGH S K. Cattle recognition: a new frontier in visual animal biometrics research[J]. Proceedings of the National Academy of Sciences India Section, 2020, 90(4): 689－708.

[8] 王静, 李蔚然, 刘业强, 等. 基于计算机视觉的养殖动物计数方法研究综述[J]. 农业机械学报, 2023, 54(增刊 1): 315－329.
WANG Jing, LI Weiran, LIU Yeqiang, et al. Review of vision counting methods and applications for farmed animals[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 1): 315－329. (in Chinese)

[9] MYAT NOE S, ZIN T T, TIN P, et al. Comparing state-of-the-art deep learning algorithms for the automated detection and tracking of black cattle[J]. Sensors, 2023, 23(1): 532.

[10] 谢秋菊, 周红, 包军, 等. 基于机器视觉的畜禽体质量评估研究进展[J]. 农业机械学报, 2022, 53(10): 1－15.
XIE Qiuju, ZHOU Hong, BAO Jun, et al. Review on machine vision-based weight assessment for livestock and poultry[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 1－15. (in Chinese)

[11] WENG Z, MENG F, LIU S, et al. Cattle face recognition based on a two-branch convolutional neural network[J]. Computers and Electronics in Agriculture, 2022, 196: 1－9.

[12] LI Z, LEI X, LIU S. A lightweigh deep learning model for cattle face recognition[J]. Computers and Electronics in Agriculture, 2022, 195: 106848.

[13] 许贝贝, 王文生, 郭雷风, 等. 基于非接触式的牛只身份识别研究进展与展望[J]. 中国农业科技导报, 2020, 22(7): 79－89.
XU Beibei, WANG Wensheng, GUO Leifeng, 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. (in Chinese)

[14] 杨蜀秦, 王振, 韩媛媛, 等. 基于融合坐标信息的改进 YOLO v4 模型识别奶牛面部[J]. 农业工程学报, 2021, 37(15): 129－135.
YANG Shuqin, WANG Zhen, HAN Yuanyuan, et al. Improved YOLO v4 model for face recognition of diary cow by fusing coordinate information[J]. Transactions of the CSAE, 2021, 37(15): 129－135. (in Chinese)

[15] WANG Y, HUA C, DING W, et al. Real-time detection of flame and smoke using an improved YOLO v4 network[J]. Signal, Image and Video Processing, 2022, 16(4): 1109－1116.

[16] LI Z, LEI X. Cattle face recognition under partial occlusion[J]. Journal of Intelligent & Fuzzy Systems, 2022, 43(1): 67－77.

[17] LI Y, ZENG J, SHAN S, et al. Occlusion aware facial expression recognition using CNN with attention mechanism[J]. IEEE Transactions on Image Processing, 2018, 28(5): 2439－2450.

[18] YUAN G, ZHENG H, DONG J. MSML: enhancing occlusion-robustness by multi-scale segmentation-based mask learning for face recognition[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2022: 3197－3205.

[19] ZENG D, VELDHUIS R, SPREEUWERS L, et al. Occlusion-invariant face recognition using simultaneous segmentation[J]. IET Biometrics, 2021, 10(6): 679－691.

[20] QIU H, GONG D, LI Z, et al. End2end occluded face recognition by masking corrupted features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(10): 6939－6952.

[21] SONG L, GONG D, LI Z, et al. Occlusion robust face recognition based on mask learning with pairwise differential siamese network[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 773－782.

[22] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// Proceedings

of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818 - 2826.
[23] GE W. Deep metric learning with hierarchical triplet loss[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 269 - 285.
[24] 赵晋陵, 詹媛媛, 王娟,等. 基于 SE - UNet 的冬小麦种植区域提取方法[J]. 农业机械学报, 2022, 53(9): 189 - 196.
ZHAO Jinling, ZHAN Yuanyuan, WANG Juan, et al. SE - UNet-based extraction of winter planting areas[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(9): 189 - 196. (in Chinese)
[25] 汪亚航,宋晓宁,吴小俊. 结合混合池化的双流人脸活体检测网络[J]. 中国图象图形学报,2020,25(7):1408 - 1420.
WANG Yahang, SONG Xiaoning, WU Xiaojun. Two-stream face spoofing detection network combined with hybrid pooling[J]. Journal of Image and Graphics, 2020,25(7):1408 - 1420. (in Chinese)
[26] JAYARAJ R, LOKESH S. Automatic image annotation using adaptive convolutional deep learning model[J]. Intelligent Automation & Soft Computing, 2023, 36(1): 481 - 497.
[27] PENG S, HUANG H, CHEN W, et al. More trainable inception-ResNet for face recognition[J]. Neurocomputing, 2020, 411: 9 - 19.
[28] MANDAL B, OKEUKWU A, THEIS Y. Masked face recognition using Resnet - 50[J]. arXiv Preprint, arXiv:2104.08997, 2021.
[29] 张宏鸣, 周利香, 李永恒, 等. 基于改进 MobileFaceNet 的羊脸识别方法[J]. 农业机械学报, 2022, 53(5): 267 - 274.
ZHANG Hongming, ZHOU Lixiang, LI Yongheng, et al. Sheep face recognition method based on improved MobileFaceNet [J]. Transactions of the Chinese Society for Agricultural Machinery,2022, 53(5): 267 - 274. (in Chinese)

(上接第 92 页)

[21] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]// 13th European Conference. Springer International Publishing, 2014: 740 - 755.
[22] LI H, KADAV A, DURDANOVIC I, et al. Pruning filters for efficient convnets[C]// 5th International Conference on Learning Representations, 2017: 149804.
[23] YANG L, XU X, ZHAO J, et al. Fusion of RetinaFace and improved FaceNet for individual cow identification in natural scenes[J/OL]. Information Processing in Agriculture,2023. http://doi.org/10.1016/j.inpa.2023.09.001.
[24] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015, 37(9): 1904 - 1916.
[25] LIU W, LIAO S, REN W, et al. High-level semantic feature detection: a new perspective for pedestrian detection[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019: 5187 - 5196.
[26] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance - IoU loss: faster and better learning for bounding box regression [C]//34th AAAI Conference on Artificial Intelligence, 2020: 12993 - 13000.
[27] 王海鹏,王正良,许威威,等. 基于三维模型的 Android 手机端人脸姿态实时估计系统[J]. 计算机应用,2015,35(8): 2321 - 2326.
WANG Haipeng,WANG Zhengliang,XU Weiwei, et al. Real-time face pose estimation system based on 3D face model on Android mobile platform[J]. Journal of Computer Applications,2015,35(8): 2321 - 2326. (in Chinese)
[28] 陈科尹,邹湘军,关卓怀,等. 基于混合蛙跳优化的采摘机器人相机标定方法[J]. 农业机械学报,2019,50(1): 23 - 34.
CHEN Keyin, ZOU Xiangjun, GUAN Zhuohuai, et al. Camera calibration method of picking robot based on shuffled frog leaping optimization[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(1): 23 - 34. (in Chinese)
[29] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition,2020: 1580 - 1589.
[30] LIU Z,LI J, SHEN Z, et al. Learning efficient convolutional networks through network slimming[C]// 16th IEEE International Conference on Computer Vision, 2017: 2736 - 2744.
[31] LEE J,PARK S,MO S,et al. Layer-adaptive sparsity for the magnitude-based pruning[J]. arXiv Preprint arXiv:2010.07611, 2020.
[32] MITTAL D, BHARDWAJ S, KHAPRA M M, et al. Recovering from random pruning: on the plasticity of deep convolutional neural networks[C]// IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018: 848 - 857.
[33] FU R, HU Q, DONG X, et al. Axiom-based grad-cam: towards accurate visualization and explanation of CNNS[J]. arXiv Preprint arXiv:2008.02312,2020.
[34] 杨慧,张思源,彭容,等. 荷斯坦犊牛反刍行为的发育规律及其与采食之间的相关性研究[J]. 动物营养学报,2022, 34(3): 1614 - 1622.
YANG Hui, ZHANG Siyuan, PENG Rong, et al. Research on developmental regularity of ruminating behavior in Holstein calves and its correlation with feed intake[J]. Chinese Journal of Animal Nutrition, 2022, 34(3): 1614 - 1622. (in Chinese)
[35] GRELET C, DRIES V V, LEBLOIS J, et al. Identification of chronic stress biomarkers in dairy cows[J]. Animal,2022, 16(5): 100502.