doi:10.6041/j.issn.1000-1298.2025.12.054

基于改进 YOLO v8 关键点检测的西门塔尔牛只个体识别方法

田振江[1]　韩　成[1]　张　超[1]　李月莹[1]　于福东[2]　钟馨慧[2]
(1.长春理工大学计算机科学技术学院,长春 130012;2.吉林省中农阳光数据有限公司,长春 130022)

摘要:随着畜牧业的智能化发展,基于计算机视觉技术的牛只个体识别在牛场养殖中扮演着越来越重要的角色。然而,在实际应用中,牛脸角度的快速变化、牛只数量的频繁变动以及模型识别速度的局限性严重影响了识别的准确性和效率。为了应对这些挑战,本文提出了一种基于改进 YOLO v8 关键点检测的西门塔尔牛个体识别方法。该方法包括 2 个阶段,在牛脸检测阶段,改进 YOLO v8 模型自动检测牛脸图像中的关键点,并进行矫正对齐和背景裁剪,使牛脸图像角度趋于一致。在牛脸识别阶段,引入了深度可分离卷积以减少参数量,利用残差连接和注意力机制增强了模型的特征提取能力,从而提升了识别准确性和速度。为验证方法的有效性,构建了包含 159 头西门塔尔牛的牛脸图像数据集。实验结果表明,本文模型在参数量和推理时间上有显著改进,单幅图像处理时间缩短至 39 ms 以内,识别准确率为 95.8%,比其他模型相比平均提高 4.8 个百分点。此外,本文方法支持实时更新牛只数据库,在增加牛脸图像后,仍能保持较高准确率。

关键词:牛脸识别;计算机视觉;YOLO v8;关键点检测
中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2025)12-0591-12　　**OSID**:

Individual Identification Method of Simmental Cattle Based on Improved YOLO v8 Keypoint Detection

TIAN Zhenjiang[1]　HAN Cheng[1]　ZHANG Chao[1]　LI Yueying[1]　YU Fudong[2]　ZHONG Xinhui[2]
(1. *School of Computer Science and Technology*, *Changchun University of Science and Technology*, *Changchun* 130012, *China*
2. *Jilin Sino Agri Sunshine Data Co.*, *Ltd.*, *Changchun* 130022, *China*)

Abstract: With the intelligent development of animal husbandry, computer-vision-based individual cattle recognition is playing an increasingly important role in modern cattle farming. However, in practical applications, the rapid changes in the angles of cattle faces, frequent fluctuations in the number of cattle, and limitations in model recognition speed severely affect the accuracy and efficiency of recognition. To address these challenges, an individual recognition method for Simmental cattle was proposed based on improved YOLO v8 keypoint detection. The method consisted of two stages: in the cattle face detection stage, the improved YOLO v8 model automatically detected keypoints in cattle face images and performed correction alignment and background cropping to make the angles of the cattle face images more consistent. In the cattle face recognition stage, depthwise separable convolution was firstly introduced to reduce the number of parameters, and then residual connections and attention mechanisms were utilized to enhance the model's feature extraction capabilities, thereby improving recognition accuracy and speed. To validate the effectiveness of the method, a dataset of cattle face images containing 159 Simmental cattle was constructed. Experimental results showed that the proposed model had significant improvements in the number of parameters and inference speed, with processing time for a single image reduced to within 39 ms and a recognition accuracy of 95.8%, an average improvement of 4.8 percentage points compared with that of other models. Additionally, the proposed method supported real-time updates to the cattle database, maintaining high accuracy even after adding new cattle face images.

Key words: cow face recognition; computer vision; YOLO v8; key point detection

收稿日期:2024-08-12　修回日期:2024-10-18
基金项目:吉林省重点科技研发项目(20180201069GX)
作者简介:田振江(1998—),男,博士生,主要从事牛脸面部识别研究,E-mail:2022101001@mails.cust.edu.cn
通信作者:韩成(1978—),男,教授,主要从事计算机视觉研究,E-mail:hancheng@cust.edu.cn

0 引言

畜牧业作为农业的重要组成部分,目前正向规模化、标准化、产业化方向发展,现代化规模养殖逐渐取代个体户散养的养殖方式[1-2]。在现代化肉牛养殖中,通过牛只个体识别技术为每头牛建立数字身份,实现精准的牛只个体追踪与信息管理,发现牛只潜在健康问题,提高养殖场养殖效率。此外,牛只个体识别技术在减少保险纠纷中发挥着关键作用,其可以为每头牛提供准确、唯一的身份标识,能够有效防止骗保和错赔现象,保障保险公司与养殖户的权益。

牛只个体身份识别方法分为接触式和非接触式 2 种:接触式识别方法包括传统标记方法和电子耳标。传统标记方法虽易于实施,但对牛只健康有影响且易被篡改;电子耳标[3]利用射频识别(Radio frequency identification,RFID)技术,识别速度快、准确性高,但受监测数量、成本和距离等限制,难以大规模推广。非接触式方法如生物特征识别和牛脸识别逐渐成熟,牛只虹膜[4]、视网膜[5]和鼻子纹理[6]等生物特征与人类眼部相似,都是独一无二的,可用于牛只个体身份识别。一些研究[7-10]通过各种策略去定位牛脸虹膜位置和虹膜的特征信息,都取得了较高的识别效果。尽管这些方法准确率高,但采集困难且不易推广。牛脸识别是一种简单、无创、易操作的方法,可以更易区分毛发纹理和面部结构特征。一些研究通过图像处理和神经网络[11]、局部二值模式纹理特征[12]和牛脸关键点[13]方法对牛脸面部特征进行识别,验证了面部特征识别方法的可行性。随着深度学习技术的发展,卷积神经网络在生物识别领域也得到应用和发展。一些研究[14-19]使用 LeNet-5、VGG、Vision-Transformer 等网络模型进行牛脸面部特征提取,实现牛脸在遮挡和脏污场景下的识别。最近一些研究聚焦多角度牛脸识别[20]和实时视频识别[21],旨在解决由于角度变化导致的牛脸特征信息不完整问题。

虽然基于深度学习技术的牛脸识别在图像识别领域表现出色,但在实际应用中,牛脸角度的快速变化、牛只数量的频繁变动(通常因购买或出售而增加或减少)以及模型识别速度的限制,往往会影响识别效果。随着牛场规模的不断扩大,需要牛脸识别技术能够快速处理这些数量变化。然而,现有的一些深度学习模型在应对牛只增减方面存在明显不足,常依赖于训练过的模型进行概率匹配。当新的牛只加入时,这些方法无法有效区分该牛只是否已存在于数据库中,导致识别的准确性和灵活性受到限制。特别是在新的牛只验证匹配时,常会出现误识别问题。

针对上述问题,本文提出一种基于改进 YOLO v8 关键点检测的西门塔尔牛只个体识别方法。在牛脸数据采集后,使用牛脸关键点检测模型,实现牛脸图像自动识别,并自动进行矫正对齐裁剪处理,减少背景干扰和角度变化带来的误差。在牛脸识别过程中,构建牛脸识别模型在降低模型参数量的同时,提升处理速度和识别准确性。此外,提出的方法能够识别和处理不在数据库中的牛只,在增加牛脸图像后,本方法仍能保持较高的准确率,提高在实际应用中的灵活性。

1 实验数据与方法

1.1 数据采集

数据于 2023 年 7—11 月在吉林省德惠牛场拍摄采集,具体拍摄牛场环境如图 1 所示。数据采集使用普通手机拍摄牛头视频,视频保存格式为 MP4,拍摄时长为 5~15 s。通过不同距离拍摄数据分析,确定最佳拍摄距离为牛脸正前方约 0.7 m,与牛头位置水平,以确保图像清晰度较高,视频拍摄角度以牛只正脸为主。由于拍摄条件限制,数据采集主要在室内进行,采集的牛只品种为西门塔尔牛,共采集 159 头牛只图像和视频数据。



(a) 牛只喂养前 (b) 牛只喂养时 (c) 牛只喂养时(室外自由活动)

图 1 牛场牛脸数据采集环境

Fig. 1 Cattle face data collection environment in cattle farms

1.2 牛脸关键点数据集

从采集的牛只中选取 85 头牛的图像数据进行手工关键点标注,其中训练集包含 60 头牛的 300 幅图像,测试集包含 25 头牛的 75 幅图像,测试集数据记为 CowK。数据标注使用 LabelMe 软件进行手动标注,标注完成后使用图像平移、增加噪点、遮挡和旋转等数据增强手段,将训练集扩充至 1 500 幅图像。标注关键点位置如图 2 所示,图 2b 标注位置为参照文献[13]中的关键点标注位置,其中 1~19 为牛脸轮廓关键点,20~23 为牛鼻关键点,24~29 为左右眼关键点。根据预实验结果,关键点数量会直接影响牛脸裁剪任务的效果,从而影响后续的牛脸识别准确性。所以本文在 29 个关键点的基础上在牛鼻、左右脸中部分割线等



图 2　牛头正脸关键点标注位置图

Fig. 2　Annotated key points on frontal cattle face diagram

重点区域添加了更多的关键点位置，增加这些区域使模型能够更好地理解牛脸的几何结构，从而提高牛脸裁剪任务的效果，最终标注如图 2c 所示的 43 个关键点。其中 1 ~ 19 为牛脸轮廓关键点，20 ~ 29 为牛鼻关键点，30 ~ 33 为左眼关键点位置，34 ~ 37 为右眼关键点位置，38 ~ 43 为牛只左右脸中部分割线。标注完成后，选择 1 幅标注好的牛脸正脸图像作为标准蒙版图像，标准蒙版图像包含 43 个关键点的标注位置，这个蒙版图像和标注位置用于后续的对齐矫正处理。

1.3　牛脸关键点检测模型

牛脸关键点的定位可以提供关于牛脸结构和形态的关键信息，因此需要 1 个能够高效、快速且精确检测牛脸关键点的模型。YOLO v8 模型虽在关键点检测任务上表现优异，但由于其模型规模较大，在资源受限设备上运行时可能会受到限制。为了提升关键点模型在这些环境中的实用性，本文对 YOLO v8 网络模型进行轻量化处理，使其在保持较高识别精度的同时，显著降低计算资源的需求。整体模型如图 3 所示，主干网络采用 MobileNetV3[22] 结构，使用轻量级的深度可分离卷积（Depthwise separable convolution，DSConv）和残差块，加速模型收敛，降低模型参数量。同时，在主干网络的最后 3 层引入卷积块注意力模块（Convolutional block attention module，CBAM）[23] 注意力机制，增强特征提取能力和模型表达能力。颈部网络采用 2 个上采样和特征融合模块组，以捕获更细节的牛脸特征。头部网络连接整个模型的输出，利用主干和颈部网络提取的特征，进行训练并做出预测。

1.4　牛脸矫正裁剪

牛脸面部裁剪矫正过程为：①使用关键点检测

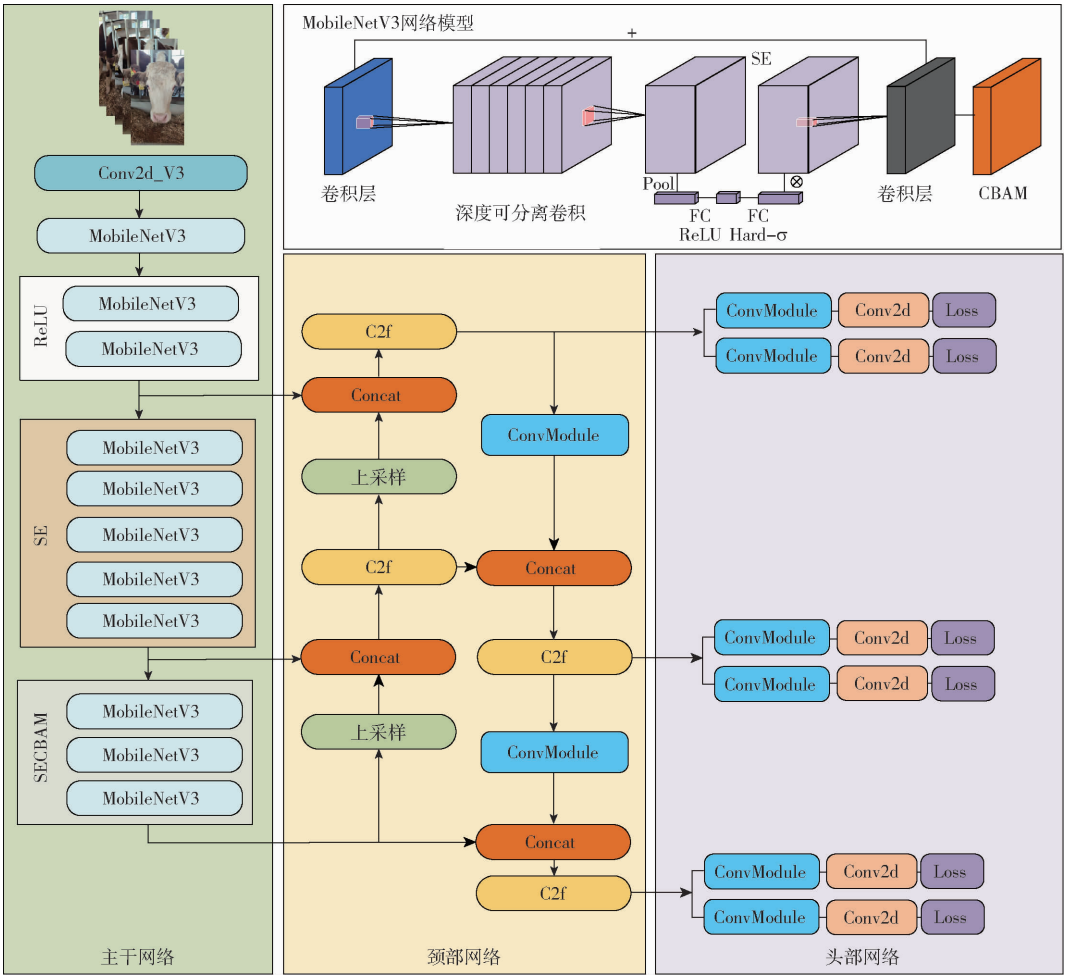

图 3　改进的 YOLO v8 模型

Fig. 3　Improved YOLO v8 model

模型对牛脸图像进行牛脸关键点检测。②将识别后的关键点与标准蒙版图像中的关键点进行对齐操作,确保两者之间位置一致。③对齐后对图像进行仿射变换,将整个牛脸图像按照对齐的比例与蒙版图像进行贴合,使牛脸各部分区域与蒙版图像区域保持一致,实现统一的牛脸形状和位置标准化,减少后续裁剪图像过程中的差异性。

在裁剪牛脸过程中,根据关键点坐标计算牛脸裁剪区域,并从原图像中裁剪出这个区域,对裁剪出的图像进行缩放,使其符合预定的标准尺寸。在缩放后的图像上,通过指定的关键点索引,连接关键点形成面部轮廓,然后提取出轮廓内的牛脸图像,处理完成后获得具有统一视角和标准尺寸的牛脸图像,为后续牛脸识别提供更清晰、更规范的输入。

1.5 数据处理

视频数据采集完成后,使用牛脸关键点检测模型对视频中的牛脸进行识别,自动将每帧识别到的牛脸图像转换为牛脸图像数据,根据检测到关键点进行裁剪对齐矫正,裁剪后尺寸为 415 像素 × 295 像素,为了确保图像质量,在裁剪过程对牛脸图像数据进行模糊度和相似度筛选,图像的模糊程度使用图像的拉普拉斯边缘强度算法评估,删除平均边缘强度小于 2 的模糊图像。对于尺寸 $x \times y$ 图像,模糊检测公式为

$$BlurScore = \frac{1}{xy}\sum_{i=0}^{x-1}\sum_{j=0}^{y-1}|\text{Laplacian}(i,j)| \quad (1)$$

式中 $BlurScore$——图像模糊分数

x、y——图像宽度、高度像素数量

i、j——图像行、列坐标位置

Laplacian(·)——图像拉普拉斯变换值[24]

牛脸相似度使用图像结构相似度(Structural similarity index measure,SSIM)[25]评估,图像 m 和图像 n 的 SSIM 公式为

$$SSIM(m,n) = \frac{(2\mu_m\mu_n + c_1)(2\sigma_{mn} + c_2)}{(\mu_m^2 + \mu_n^2 + c_1)(\sigma_m^2\sigma_n^2 + c_2)} \quad (2)$$

式中 μ_m、μ_n——图像 m、n 像素均值

σ_m^2、σ_n^2——图像 m、n 像素方差

σ_{mn}——图像 m、n 协方差

c_1、c_2——常数,防止分母为 0

为确保同一头牛有更高的区分度,本文对相似度大于 0.5 的图像进行删除,确保所有图像两者之间的平均相似度在 0.5 以下。在进行模糊度和相似度筛选后,对于牛脸图像数量不足 40 幅的牛只,进行删除并不再用于后续处理,确保数据集的质量和模型训练的有效性。

根据关键点检测模型分别制作了以下 3 个图像数据集:29 个关键点裁剪图像数据集,通过使用 29 个关键点检测模型进行关键点检测,提取牛脸关键点轮廓区域内图像;43 个关键点裁剪图像数据集,通过使用 43 个关键点检测模型进行关键点检测,提取牛脸关键点轮廓区域内图像;无裁剪图像数据集,仅通过 43 个关键点检测算法对牛脸位置进行识别,保存牛脸识别框内图像。3 个数据集图像如图 4 所示,其中原始图像尺寸为 1 280 像素 ×760 像素,无裁剪图像尺寸为 415 像素 ×415 像素,29 个关键点和 43 个关键点裁剪图像尺寸为 415 像素 ×295 像素。

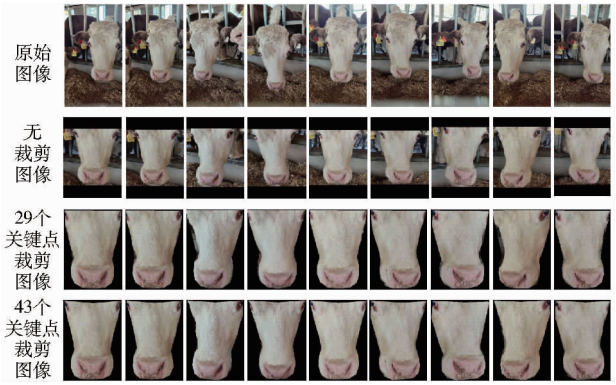

图 4 牛脸裁剪数据集

Fig. 4 Cattle face cropping dataset

1.6 牛脸识别数据集划分

为了验证本文提出的牛脸识别方法,将数据集划分为原有牛只数据集和新入库牛只数据集,原有牛只数据集共 109 头牛,每头牛只分别选择 80 幅图像,按比例 14∶3∶3进行划分,其中 56 幅用于训练牛脸面部特征模型,12 幅图像用于将牛只注册到数据库,12 幅图像用于验证匹配数据库。对训练集进行数据增强操作,最终训练图像数量共 30 520 幅,记为 CowTr;注册图像数量共 1 308 幅,记为 CowTre;测试图像数据共 1 308 幅,记为 CowTc。测试图像与注册图像进行匹配评估模型识别效果。新入库牛只数据集一共 50 头,每头牛只提取 40 幅图像,然后从中选择 12 幅图像与注册的牛只进行匹配。测试图像共 600 幅,记为 CowF,评估牛脸识别模型在验证新牛只图像时的准确性和可靠性。

2 牛脸识别模型架构

牛脸识别技术可定义为输入牛脸图像,提取牛脸特征,并与数据库中的已知牛脸特征进行匹配。在第 1 阶段,本文方法可以自动检测牛脸图像中的关键点位置,进行矫正对齐,裁剪背景区域,从而降低与数据库牛脸图像的差异。在第 2 阶段,使用 FaceNet[26]框架进行牛脸识别任务。输入图像使用

第 1 阶段处理后的牛脸裁剪图像，以特殊的三元组数据导入模式作为牛脸特征提取网络输入，使同一身份的牛脸特征在特征空间中更加接近，而不同身份的牛脸特征更加分散。牛脸图像经特征提取网络处理后，对提取的特征进行 L2 归一化，最终获得 128 维牛脸特征向量。最后，使用牛脸特征提取模型对注册集和测试集进行特征编码，完成牛脸数据的注册。然后通过计算测试集编码与注册集编码之间的欧氏距离，进行特征匹配并输出识别结果。整个牛脸识别任务框架如图 5 所示。

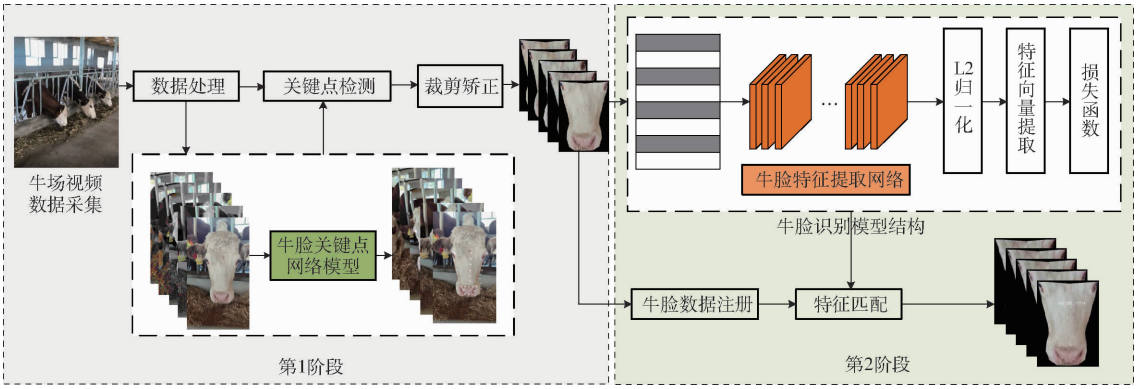

图 5　牛脸识别框架

Fig. 5　Framework for cattle face recognition

2.1　牛脸特征提取网络

牛脸特征提取网络使用改进的 Inception－ResNetv2[27] 网络结构，为提高模型识别速度，引入 Inception 架构、残差网络和 DSConv。

将 Inception 模块的输出添加到输入中，统一输入与输出维度，便于进行残差加法运算。由于裁剪后牛脸图像不包含背景，需要对牛脸的细节特征进行精准捕捉。为此，在 Inception－resnet－A、B、C 3 个模块后加入 CBAM 注意力机制，以增强特征提取能力。改进 Inception－ResNetv2 网络包括 6 个模块：Stem、Inception－resnet－A、Reduction－A、Inception－resnet－B、Reduction－B 和 Inception－resnet－C。Stem 模块负责对输入进行初步处理和特征提取；Inception－resnet－A、B、C 模块分别在不同层次上提取特征；Reduction－A 和 Reduction－B 模块用于缩减特征图尺寸，从而减少计算量和参数量。这种设计可有效减少模型参数量，加速模型训练和推理，同时在特征提取能力和模型识别性能上有显著提升。网络结构如图 6 所示。

2.2　损失函数

在牛脸识别方法中，损失函数由三元组损失和交叉熵损失共同组成。通过结合这两种损失，模型不仅能够有效地学习样本间的相对距离，还能进行精确的分类，三元组损失[28]是一种用于训练人脸识

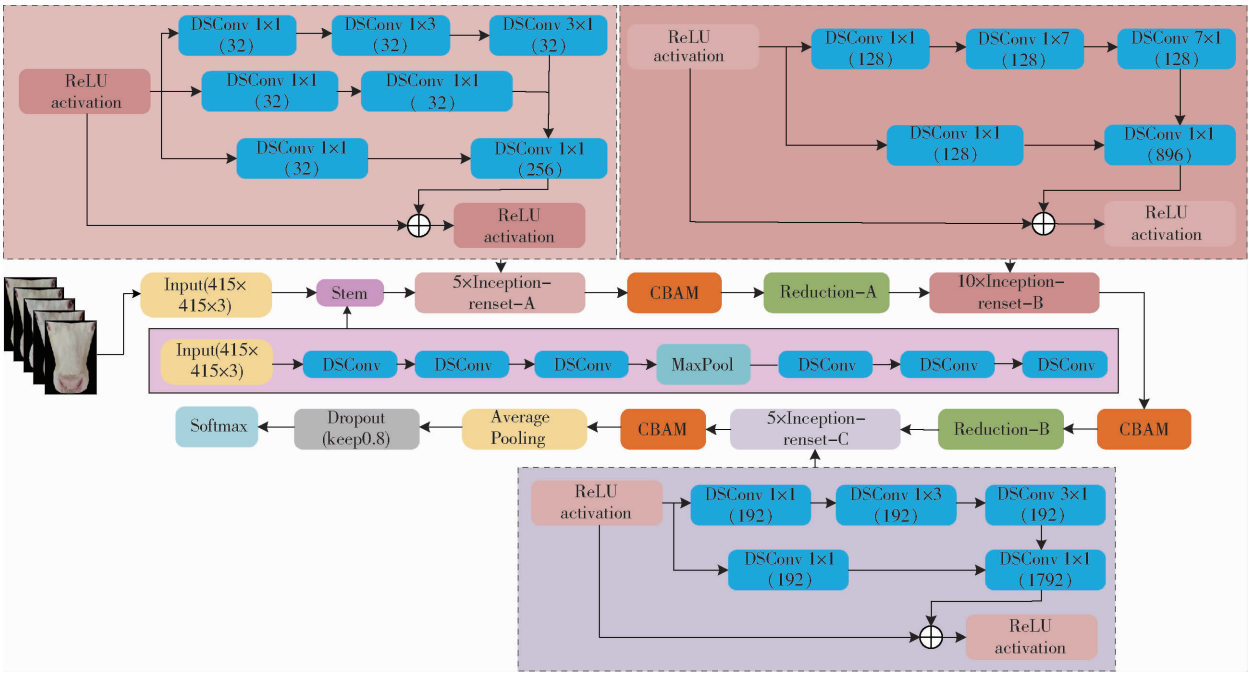

图 6　改进 Inception－ResNet 网络结构

Fig. 6　Improved Inception－ResNet network architecture

别和其他识别学习任务的常用损失函数。该函数基于三元组样本,对于牛脸数据集中,在每个样本设置锚点牛脸图像、正例牛脸图像和负例牛脸图像,正例图像与锚点图像同一身份,而负例图像与锚点图像不同身份。整体损失函数的目标是使得锚点图像与其正例图像之间的距离缩小,与负例图像之间的距离增加,三元组损失定义为

$$L_{triplet}=\sum_{k=1}^{o}\left[\|h(r_k^a)-h(r_k^t)\|_2^2-\|h(r_k^a)-h(r_k^s)\|_2^2+\alpha\right]_+ \tag{3}$$

式中 o——训练样本数量

r_k^a——第 k 个样本锚点图像

r_k^t、r_k^s——第 k 个样本正例、负例图像

$\|\cdot\|_2^2$——欧几里得距离

$h(\cdot)$——提取特征的函数

α——间隔的超参数,规定正图像与负图像之间的最小距离

$[\cdot]_+$——取正值函数,即当括号内值为负数时取 0,否则取该值本身

这样的训练目标有助于网络能够将同一牛脸图像之间的特征向量尽可能靠近,不同牛脸图像之间的特征向量尽可能远离,从而实现更好的牛脸验证性能,三元组损失示意图如图 7 所示。

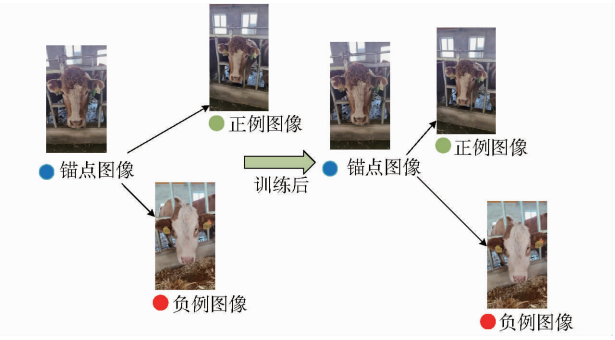

图 7 三元组损失示意图

Fig. 7 Triplet loss schematic

交叉熵损失用于分类任务,通过最小化预测概率分布与真实分布之间的交叉熵来训练模型。模型在输出层使用 Softmax 函数,计算每个类概率,并通过负对数似然损失来计算损失值。

2.3 实验环境与评价指标

实验环境:CPU 配置为 13th Gen Intel(R) Core(TM) i9-13900HX,GPU 为 GeForce RTX4080,具备 12 GB 显存和 32 GB 内存,操作系统为 Windows 11 系统。采用 Pytorch 深度学习框架,对模型训练学习率动态调整设值,范围为 0.001 ~ 0.000 01,批量大小为 12,迭代次数为 100。采用自适应矩阵估计算法对模型进行优化,损失值计算则采用三元组损失函数和交叉熵损失结合方法。

在牛脸关键点检测阶段,使用归一化均值误差(Normalized mean error,NME)算法计算牛脸关键点坐标与真实牛脸坐标之间的误差,衡量关键点检测算法准确度。

使用牛脸左右眼距(眼中心距离)作为归一化因子。在牛脸识别阶段时采取准确率(Accuracy)、精确率(Precision)、召回率(Recall)、假正率(FPR)和 F1 分数(F1 score)对本文模型进行对比验证。

为了全面评估模型的推理性能,本文统计模型的参数量和推理时间。其中,推理时间指模型在单次推理时消耗的时间,而平均推理时间则为多次推理后平均消耗时间,用于衡量模型实时性表现。

使用目标关键点相似度(Object keypoint similarity,OKS)作为关键点检测精度的评价指标[29]。OKS 能够衡量检测到的关键点与真实关键点之间的相似度,并基于不同阈值计算精确率、召回率和平均精度,从而全面评估模型在关键点检测任务中的表现。

在实验部分除 3.3 节使用 29 个关键点裁剪图像数据集和无裁剪图像数据集外,其余实验部分均使用 43 个关键点裁剪图像数据集。

3 结果与分析

3.1 关键点检测对比结果

在牛脸关键点检测过程中,采用 NME 算法评估模型性能,其中归一化因子 d 选取牛脸左右眼 4 个关键点中心点之间的距离,避免不同牛脸大小和姿态对结果的干扰,确保评估一致性和可靠性。对比 YOLO v8 模型与本文模型在牛脸正面关键点识别任务中的表现,如表 1 所示。

表 1 牛脸关键点检测对比结果

Tab. 1 Comparison of cattle face keypoint detection results

模型	数据集	NME/%	推理时间/ms	参数量	牛脸识别任务准确率/%
YOLO v8	CowK	3.11	25	4.03×10^6	95.9
本文	CowK	3.52	10	1.95×10^6	95.8

为了验证 NME 指标对后续牛脸识别精度的影响,在相同的 43 个关键点基础下,对基于 YOLO v8 模型的裁剪数据和基于本文模型的裁剪数据分别进行了实验。实验结果表明,在牛脸关键点检测任务中,YOLO v8 与本模型 NME 差值 0.41 个百分点,对应的识别准确率相差约 0.1 个百分点。但是本文方法推理时间 10 ms 与改进前相比提升 60%,参数量下降 51.6%。因此,虽然原模型在精度上略有优势,但改进模型在计算效率和资源占用上显著优于

原模型，适合实际应用场景中对实时性和设备资源要求较低的任务。

3.2　关键点裁剪矫正技术评估

为了全面评估本文提出的关键点裁剪矫正技术对牛脸识别结果的影响，分别对 29 个关键点裁剪图像数据集、43 个关键点裁剪图像数据集以及无裁剪图像数据集进行了训练实验。在实验评估过程，除数据集类型不同外，其余实验条件均保持一致，数据划分比例也与 43 个关键点裁剪数据集的划分比例一致，以确保结果的公平性与可比性。通过不同裁剪方案的对比分析，可以更准确地了解关键点数量和裁剪策略对模型性能的影响，结果见表 2。

表 2　关键点裁剪矫正技术下的牛脸识别对比结果

Tab. 2　Comparison results of cattle face recognition under keypoint cropping and correction techniques

%

类型	评估数据集	准确率	精确率	召回率	假正率	F1 分数
未使用裁剪	CowTc + CowF	79.5	80.2	91.8	45.2	85.6
29 个关键点裁剪	CowTc + CowF	87.9	88.3	93.9	23.7	91.0
43 个关键点裁剪	CowTc + CowF	95.8	97.2	96.5	5.7	96.8

由表 2 可知，未使用裁剪技术的模型在处理牛脸图像时，受图像背景和其他干扰因素的影响较大，各指标均最低。使用 29 个关键点进行裁剪后，模型能够更好地聚焦于牛脸的关键特征，减少了背景干扰，显著提高了整体各指标。使用 43 个关键点进行裁剪后，模型能够更加精细地提取牛脸边缘的特征信息，模型的表现进一步提升，达到最佳效果。综上所述，裁剪矫正技术在牛只面部个体识别模型的准确性和稳健性具有积极作用，而且关键点数量也会影响最终的识别效果。

3.3　不同网络模型牛脸识别对比结果

在图像经过 43 个关键点模型矫正裁剪处理后，所有牛脸的姿态和大小均与标准牛脸相似，整体特征相似性较高。本文选择匹配欧氏距离阈值为 0.6 进行特征提取和验证匹配，在相同实验环境和参数设置下，将 FaceNet 结构中主干网络分别替换为 MobileNetv2[30]、Inception - ResNetv2[27]、VGG1[31]、Vision - Transformer[32]、Swin - Transformer[33] 和 ResNet50[34] 网络进行对比，不同网络模型对比结果如表 3 所示。由表 3 可知，本文模型较其他模型的牛脸识别准确率、召回率和 F1 分数均更优，本文模型的精确率和假正率没有取得最优结果，经过分析发现 Swin - Transfomer 模型的精确率和假正率都高于本模型，尽管 Swin - Transformer 在指标上表现较好，但其整体准确率、召回率、F1 分数仍不及本文模型，且该模型的推理时间较慢，不适合于实时牛脸识别任务，本文模型在识别牛脸时更加全面稳定。

表 3　不同网络模型牛脸识别对比结果

Tab. 3　Comparison results of cattle face recognition with different network models

%

网络模型	评估数据集	准确率	精确率	召回率	假正率	F1 分数
MobileNetv2	CowTc + CowF	91.3	94.9	91.9	10.0	93.4
Inception - ResNetv2	CowTc + CowF	91.8	92.8	95.1	14.8	93.9
VGG16	CowTc + CowF	86.7	89.9	90.2	20.5	90.0
Vision - Transfomer	CowTc + CowF	90.7	94.0	92.0	11.8	93.0
Swin - Transfomer	CowTc + CowF	94.0	98.4	92.5	3.0	95.4
ResNet50	CowTc + CowF	91.1	95.1	91.3	9.5	93.2
本文模型	CowTc + CowF	95.8	97.2	96.5	5.7	96.8

3.4　添加新增牛只数据实验结果

在本实验中，对数据库中添加不同规模的牛只数据对模型准确率的影响进行评估，如表 4 所示，即在数据库中分别添加 50 头牛的 50 幅图像、100 幅图像、250 幅图像、350 幅图像和 600 幅图像，以进行对比验证。测试数据集使用数据库中的牛只验证集 CowTc，与之进行匹配数据使用数据库中的牛只注册集 CowTre 和新添加的牛只图像。通过不同规模批次的实验，分析数据规模对各网络模型识别准确率的影响。

由表 4 可知，不同模型在增加图像数量后的表现存在显著差异，MobileNetv2、VGG16、ResNet50 模型在增加图像数量后，准确率基本保持不变，可能是由于这些模型对数据集变化具有较高的鲁棒性，能够在不同数量图像上保持稳定性能，但这些模型的准确率相对较低。而 Inception - ResNetv2、Vision -Transformer、Swin - Transformer 模型和本文模型，在增加图像数量后，模型对数据集变化较为敏感，准确率有所下降。本文模型从 96.4% 下降到 95.3%，始终高于其他模型准确率。这表明模型在处理一定规模的数据变化时仍保持了较高准确率，但也反映了模型在应对更多样化数据时可能面临一

表 4 牛脸识别在添加新增牛只数据下准确率对比

Tab. 4 Comparison of accuracy of cow face recognition with addition of new cattle data %

网络模型	测试数据集	原模型准确率	50 幅图像	100 幅图像	250 幅图像	350 幅图像	600 幅图像
MobileNetv2	CowTc	91.9	91.9	91.7	91.7	91.7	91.7
Inception - ResNetv2	CowTc	94.4	94.2	94.1	94.0	93.9	93.8
VGG16	CowTc	89.5	89.5	89.5	89.5	89.5	89.5
Vision - Transfomer	CowTc	91.8	91.5	91.5	91.4	91.2	91.1
Swin - Transfomer	CowTc	92.4	92.4	92.1	91.8	91.7	91.5
ResNet50	CowTc	90.9	90.9	90.9	90.9	90.9	90.9
本文模型	CowTc	96.4	96.2	95.8	95.7	95.4	95.3

定挑战,如在实际应用中加入大规模数据,可能需要对新加入的牛脸图像进行训练处理以提升牛脸识别模型准确率。

3.5 消融实验

为验证本文模型引入的 DSConv 和 CBAM 注意力机制模块对牛脸识别模型的影响,本文在与 3.3 节相同的实验环境和参数设置下进行消融实验,结果如表 5 所示。

相较于 Inception - ResNetv2(基础模型),将 DSConv 引入到模型中,模型参数量大幅减少,同时识别速度提升,表明 DSConv 的引入有效降低了模型计算复杂度,提高了模型运行效率,但模型精确率和假正率有所下降,反映出在降低计算成本的同时,模型的识别精度也受到了一定影响。通过将 CBAM 注意力机制加入到 Inception - ResNetv2 + DSConv 模型后,模型整体性能得到显著提升。所有指标均达到最高值,虽然模型参数量有所上升,但仍优于 Inception - ResNetv2 模型,总体表明引入 DSConv 和 CBAM 注意力机制适合应用于对识别精度要求较高的场景。

表 5 牛脸识别模型消融实验对比结果

Tab. 5 Comparison results of ablation experiments for cattle face recognition models

Inception - ResNetv2	DSConv	CBAM	准确率/%	精确率/%	召回率/%	假正率/%	F1 分数/%	参数量	推理时间/ms
√	×	×	90.9	92.1	94.5	16.3	93.3	2.26×10^7	57
√	√	×	91.0	91.3	95.6	18.5	93.4	1.54×10^7	35
√	√	√	95.8	97.2	96.5	5.7	96.8	1.95×10^7	39

注:√表示采用此模块,×表示未采用此模块。

3.6 牛脸识别性能实验结果

整个牛脸识别过程推理时间和参数量如表 6 所示。由表 6 可知,本文模型可显著减少总推理时间,与任意牛脸特征模型组合,总推理时间均明显低于使用 YOLO v8 的组合。虽然 MobileNetv2 模型推理时间低于本文模型,但是模型的准确率明显低于本文模型。尽管 Vision - Transformer 和 Swin - Transformer 在牛脸识别任务中有较高的识别精度,

表 6 牛脸识别性能评估

Tab. 6 Cattle face recognition performance evaluation

牛脸检测模型	参数量	检测推理时间/ms	牛脸识别模型	参数量	识别推理时间/ms	总计推理时间/ms
YOLO v8	4.03×10^6	25	MobileNetv2	3.21×10^6	22	47
			Inception - ResNetv2	3.04×10^7	35	60
			VGG16	1.47×10^7	39	64
			Vision - Transfomer	8.56×10^7	43	68
			Swin - Transfomer	2.75×10^7	135	160
			ResNet50	2.35×10^7	48	73
			本文模型	1.95×10^7	29	54
本文模型	1.95×10^6	10	MobileNetv2	3.21×10^6	22	32
			Inception - ResNetv2	3.04×10^7	35	45
			VGG16	1.47×10^7	39	49
			Vision - Transfomer	8.56×10^7	43	53
			Swin - Transfomer	2.75×10^7	135	145
			ResNet50	2.35×10^7	48	58
			本文模型	1.95×10^7	29	39

但其总推理时间较长，不适用于实时性要求较高的应用场景。综合以上评估结果，本文模型在牛脸识别任务中，整体表现更优。

4　讨论

4.1　关键点检测准确度对牛脸姿态对齐效果

牛脸图像易受图像质量、栏杆遮挡和牛脸姿态等外界因素影响，导致关键点检测不准确或缺失，从而影响姿态对齐的效果。本文模型在采集牛脸视频数据时，会对拍摄位置和角度提出限制，影响获取较高质量的牛脸图像。

采用改进 YOLO v8 关键点检测模型，在关键点检测模型训练中采用局部遮挡数据增强方法，增强模型在遮挡情况下检测鲁棒性。此外，在检测过程如果牛脸关键特征足够清晰，即使存在部分遮挡，该模型仍能准确识别所有关键点。然而，当牛脸姿态偏离正面过大或牛脸特征被大范围遮挡时，模型的关键点检测会整体失败，而非仅限于部分关键点的缺失。因此在处理牛脸视频数据时，关键点检测模型会自动删除无法识别的牛脸图像。这种筛选机制有助于提高后续姿态对齐的质量，但也可能导致可用的训练数据量减少，影响模型的训练效果和泛化能力。未来研究中，将进一步探索多角度数据采集及处理方法，优化模型算法，力求在去除低质量数据的同时，最大限度提高在不同姿态和遮挡情况下的检测精度，保留更多有用的训练样本。

为了验证关键点检测模型的准确性，在数据采集中采集部分受遮挡的牛脸图像，通过分析这些遮挡数据特点，使用牛脸关键点测试集图像，构建牛脸遮挡数据集，如图 8 所示。该数据集模拟不同遮挡位置、形状和范围，遮挡比例为 5% ~50%。

在不同遮挡条件下对模型关键点检测平均精度进行评估，结果如表 7 所示。随着遮挡比例增加，模型关键点检测精度逐渐下降。特别是在牛只面部内部单个遮挡和多个遮挡情况下，平均精度在遮挡区域不断扩大情况下逐渐降低。在 50% 遮挡时降低至 50% 以下，且有超过 1/3 的牛只在该条件下无法正确识别。相比之下，当遮挡发生在牛只面部外部区域时，模型表现相对稳定，平均精度在不同遮挡比例下均保持较高水平，说明牛脸面部区域外的遮挡对模型检测影响较小。

结合实验结果可知，大面积遮挡会导致关键点不准确和无法识别的问题，影响牛脸姿态对齐效果。因此，未来研究将重点关注对遮挡区域和缺损区域的处理，进一步提高模型在复杂环境下的检测准确性和姿态对齐效果。



图 8　实际牛脸遮挡图像与牛脸遮挡数据集

Fig. 8　Actual cattle face occlusion image with cow face occlusion dataset

表 7　不同遮挡情况下牛脸关键点平均精度

Tab. 7　Average precision of cattle face keypoint detection under different occlusions　%

牛脸遮挡类别	遮挡比例/%						
	0	5	10	15	20	25	50
内部单个区域遮挡	95.7	94.3	91.4	90.9	85.7	81.5	41.7
内部多个区域遮挡	95.7	94.3	92.9	91.4	85.1	83.7	47.8
外部区域遮挡	95.7	95.7	95.7	95.7	94.3	92.7	82.9

4.2　方法适用性与局限性

在牛只个体识别研究中，本文模型对拍摄位置提出了限制条件。在此条件下采集牛脸图像具有较高的质量和显著的特征，有助于提升关键点检测的准确性。通过精确裁剪、对齐和矫正，减少图像间差异和背景干扰。结合 FaceNet 进行特征提取和识别，在西门塔尔牛品种个体识别中，本文模型识别准确率达到 95.8%，且单幅图像处理时间缩短至 39 ms 以内，满足快速处理需求。

为进一步验证本文模型在不同场景和不同牛种中的适用性，对多种场景与不同品种牛只进行实验，结果如图 9 所示。

从图 9 可知，本文模型在牛只面部图像质量较高的情况下具备较好的识别表现。然而，由于本文所使用数据集牛只品种为西门塔尔牛，且训练数据主要集中于正脸角度，因此在面对牛种差异较大或面部角度显著偏离正面的情况下，本文模型无法实现牛脸面部位置定位和关键点检测。此外，本文模

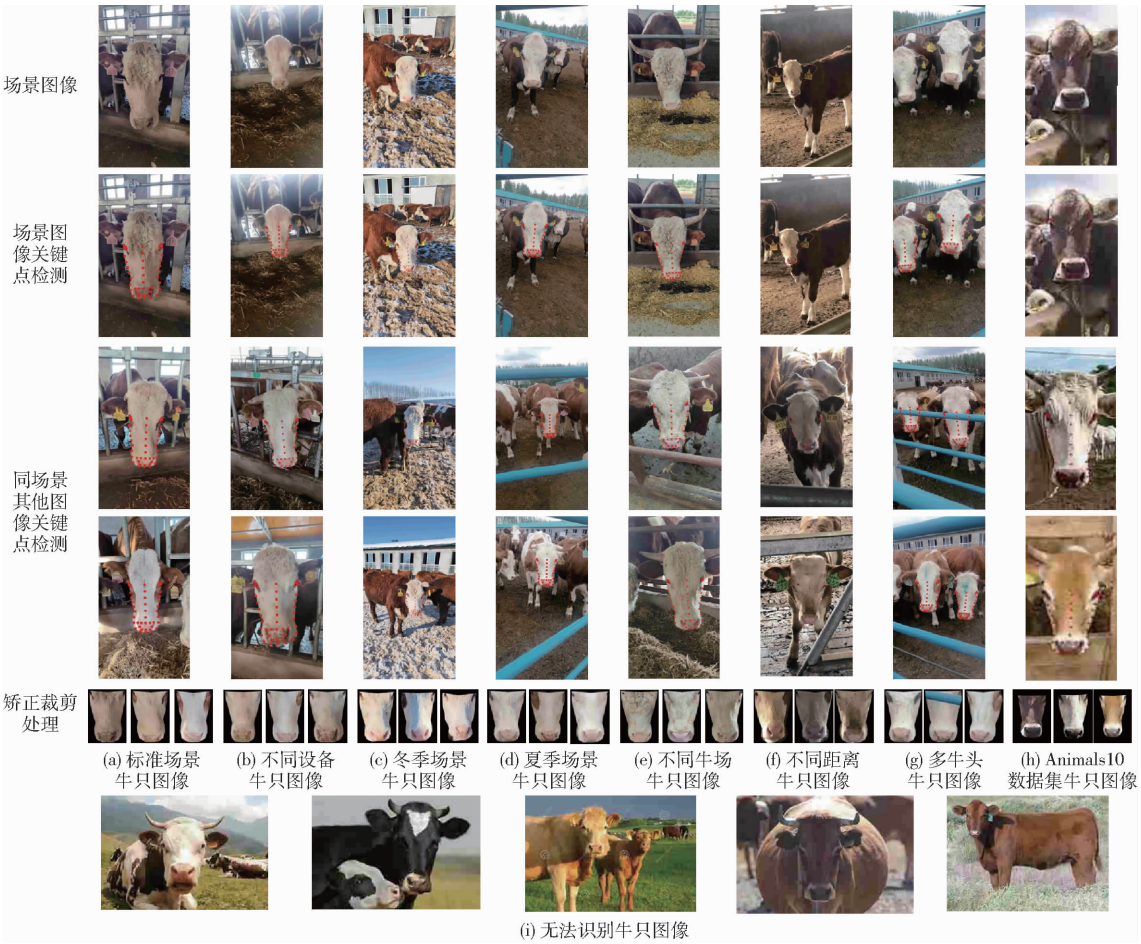

图 9 不同场景和不同品种下的识别效果

Fig. 9 Recognition performance in different environments and for various cattle breeds

型在复杂场景下仍存在一定局限性,比如,当牛只面部被大面积遮挡或光照条件极端时(过亮或过暗),关键点检测的准确性会大幅下降,导致姿态对齐效果不佳,进而影响最终的识别结果。

5 结论

(1)提出了基于改进 YOLO v8 关键点检测的西门塔尔牛只个体识别方法,通过移动手机设备采集真实牛场环境视频图像,利用改进 YOLO v8 模型关键点检测技术自动识别牛脸关键点位置,对关键点区域进行矫正对齐和背景裁剪处理,对处理后的数据进行数据增强操作,制作了牛脸图像数据集。

(2)提出的关键点检测模型 NME 值为 3.52%,误差较低。与原 YOLO v8 模型相比虽有下降,但对牛脸识别影响不大,且改进模型的参数量和推理时间都有明显提升。本文提出的牛脸识别模型结合残差网络、深度可分离卷积和注意力机制,实现了对牛只个体身份的准确识别。与其他同类模型相比,本文模型识别准确率达到 95.8%,证明了在牛脸识别任务中的有效性。

(3)牛脸识别性能实验结果表明,在牛脸图像数量变化的情况下,仍有较高的准确率,并可使单幅图像处理时间缩短到 39 ms 以内,有助于模型在移动端等计算量小的设备上应用。

参 考 文 献

[1] 张姝瑾,许兴时,邓洪兴,等. 基于 YOLO v8n-seg-FCA-BiFPN 的奶牛身体分割方法[J]. 农业机械学报,2024,55(3):282 - 289,391.
ZHANG Shujin, XU Xingshi, DENG Hongxing, et al. Segmentation model of cow body parts based on YOLO v8n-seg-FCA-BiFPN[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(3):282 - 289,391. (in Chinese)

[2] GUO Y, HONG W, WU J, et al. Vision-based cow tracking and feeding monitoring for autonomous livestock farming: the YOLOv5s-CA + DeepSORT-Vision Transformer[J]. IEEE Robotics & Automation Magazine, 2023,30(4):68 - 76.

[3] 董正奇. 基于体温监测和饮水行为记录的牛只监管耳标设计[D]. 杨凌:西北农林科技大学,2024.
DONG Zhengqi. Cow regulatory ear tag design based on body temperature monitoring and drinking behavior records[D].

Yangling: Northwest A&F University, 2024. (in Chinese)
[4] 司永胜,王朝阳,张艳,等.基于深度与传统特征融合的非限制条件下奶牛个体识别[J].农业机械学报,2023,54(6):272-279.
SI Yongsheng, WANG Zhaoyang, ZHANG Yan, et al. Individual identification of dairy cows under unrestricted conditions based on fusion of deep and traditional features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 272-279. (in Chinese)
[5] LOPEZ R B. Diseases of the heart and blood vessels and their relation to the eye[J]. Clinical and Experimental Optometry, 1943, 26(2): 60-69.
[6] KUMAR S, SINGH S K, ABIDI A I, et al. Group sparse representation approach for recognition of cattle on muzzle point images[J]. International Journal of Parallel Programming, 2018, 46(5): 812-837.
[7] ALLEN A, GOLDEN B, TAYLOR M, et al. Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland[J]. Livestock Science, 2008, 116(1-3): 42-52.
[8] 孔强, 赵林度. 虹膜识别在肉类食品安全追溯系统中的应用及关键技术研究[J]. 中国安全科学学报,2009,19(3):155-160.
KONG Qiang, ZHAO Lindu. Research on the application of iris recognition in meat-food supply chain management[J]. China Safety Science Journal, 2009, 19(3): 155-160. (in Chinese)
[9] 盛大玮. 牛眼虹膜识别技术研究[D]. 上海:华东师范大学,2010.
SHENG Dawei. Research on technology of cattle's iris recognition[D]. Shanghai: East China Normal University, 2010. (in Chinese)
[10] 李超, 赵林度. 牛眼虹膜定位算法研究及其在肉食品追溯系统中的应用[J]. 中国安全科学学报,2011,21(3):124-130.
LI Chao, ZHAO Lindu. Research on cattle iris localization algorithm and its application in meat food tracking and traceability system[J]. China Safety Science Journal, 2011, 21(3): 124-130. (in Chinese)
[11] KIM H T, IKEDA Y, CHOI H L. The identification of Japanese black cattle by their faces[J]. Asian-Australasian Journal of Animal Sciences, 2005, 18(6): 868-872.
[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] 宋肖肖.牛脸特征点检测的研究与实现[D]. 杨凌:西北农林科技大学,2017.
SONG Xiaoxiao. Research and implementation of cattle face feature points detection[D]. Yangling: Northwest A&F University, 2017. (in Chinese)
[14] 孙曦, 赵平. 基于卷积神经网络的牛脸识别方法研究[J]. 大众标准化,2020(5):59-60.
[15] 黄俊华, 田壮. 基于改进 VGG 卷积神经网络的奶牛牛脸识别研究[J]. 湖北农机化,2019(13):126.
[16] 郑鹏,沈雷,刘浩,等.基于 Vision Transformer 模型的牛脸识别算法[J]. 杭州电子科技大学学报(自然科学版),2022,42(6):40-46.
[17] 刘世锋,常蕊,李斌,等.基于脸部 RGB-D 图像的牛只个体识别方法[J]. 农业机械学报,2023,54(增刊 1):260-266
LIU Shifeng, CHANG Rui, LI Bin, et al. Individual identification of cattle based on RGB-D images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 1): 260-266. (in Chinese)
[18] 杨胜楠,赵建敏,杨梅,等.基于去偏置项 SoftMax 和紧致度量损失函数的牛脸识别方法[J]. 黑龙江畜牧兽医,2024(4):36-42.
YANG Shengnan, ZHAO Jianmin, YANG Mei, 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. (in Chinese)
[19] 叶孟珂,李宝山,杨梅,等.基于深度学习的牛脸识别系统设计[J]. 黑龙江畜牧兽医,2024(4):43-48.
YE Mengke, LI Baoshan, YANG Mei, et al. Design of cattle face recognition system based on deep learning[J]. Heilongjiang Animal Science and Veterinary Medicine, 2024(4): 43-48. (in Chinese)
[20] 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: 106871.
[21] BERGMA N, YITZHAKY Y, HALACHMI I. Biometric identification of dairy cows via real-time facial recognition[J]. Animal, 2024, 18(3): 101079.
[22] 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.
[23] WOOS, PAPK J, LEE J Y, et al. Cbam: convolutional block attention module[C]// Proceedings of European Conference on Computer Vision, 2018: 3-19.
[24] PRETUZ S, PUIG D, GARCIA M A. Analysis of focus measure operators for shape-from-focus[J]. Pattern Recognition, 2013, 46(5): 1415-1432.
[25] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
[26] 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.
[27] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on

learning[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31(1):4278 - 4284.
[28] BREDIN H. TristouNet: triplet loss for speaker turn embedding[C]// Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). IEEE,2017:5430 - 5434
[29] 宋怀波,华志新,马宝玲,等. 基于 SimCC-ShuffleNetV2 的轻量化奶牛关键点检测方法[J]. 农业机械学报,2023,54(10):275 - 281,363.
SONG Huaibo, HUA Zhixin, MA Baoling, et al. Lightweight keypoint detection method of dairy cow based on SimCC-ShuffleNetV2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 275 - 281, 363. (in Chinese)
[30] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510 - 4520.
[31] LIU S, DENG W. Very deep convolutional neuralnetwork based image classification using small training sample size[C]// 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). IEEE, 2015: 730 - 734.
[32] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022,45(1):87 - 110.
[33] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021:10012 - 10022.
[34] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.

(上接第 590 页)

[26] ZHANG H, WU C, ZHANG Z, et al. Resnest: split-attention networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2736 - 2746.
[27] DAI Y, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3560 - 3569.
[28] CHANG D, DING Y, XIE J, et al. The devil is in the channels: mutual-channel loss for fine-grained image classification[J]. IEEE Transactions on Image Processing, 2020, 29: 4683 - 4695.
[29] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision,2017: 618 - 626.
[30] 冯璐,李奇峰,丁露雨,等. 自然通风牛舍环境监测传感器布局策略对比研究[J]. 家畜生态学报,2022,43(2):74 - 81.
FENG Lu, LI Qifeng, DING Luyu, et al. Comparative research on sensor layout strategy of natural ventilation cowshed environment monitoring[J]. Journal of Livestock Ecology, 2022, 43(2): 74 - 81. (in Chinese)
[31] 刘爱瑜,冯曼,李永亮,等. 卷帘舍环境参数与奶牛泌乳性能及行为的相关性分析[J]. 畜牧兽医学报,2023,54(2):617 - 629.
LIU Aiyu, FENG Man, LI Yongliang, et al. Correlation analysis between environmental parameters of rolling shed and lactation performance and behavior of dairy cow[J]. Acta Veterinaria et Zootechnica Sinica, 2023, 54(2): 617 - 629. (in Chinese)
[32] ANDREW W, GAO J, MULLAN S, et al. Visual identification of individual Holstein-Friesian cattle via deep metric learning [J]. Computers and Electronics in Agriculture, 2021, 185: 106133.
[33] WANG Y, XU X, WANG Z, et al. ShuffleNet - Triplet: a lightweight RE-identification network for dairy cows in natural scenes[J]. Computers and Electronics in Agriculture, 2023, 205: 107632.