

基于逐元素动态融合与自适应损失函数的牛脸识别算法

齐咏生^{1,2} 赵锦程¹ 刘利强^{1,2} 李永亭^{1,2} 王朝霞^{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: 针对牛只识别任务中存在特征学习敏感性差、牧场背景复杂、冗余信息干扰严重及高相似度个体区分困难的问题, 提出一种基于逐元素动态融合与自适应损失函数相结合的牛只面部识别方法。采用 StarNet 架构对特征提取网络进行重新设计, 构建一种融合多层次特征的高效网络, 用于提炼出关键的视觉信息, 增强模型对细微差异的敏感性; 提出一种逐元素特征融合模块, 对低维特征进行筛选, 对不同特征图进行加权融合, 确保重要特征得到保留, 而无关或冗余的特征则被抑制和去除; 设计一种动态自适应损失函数, 将 ArcFace Loss 的角度边界进行自适应调整, 从而增强模型对高相似度样本的区分能力, 且能够自适应地平衡不同类别之间的特征分布, 显著提升整体识别性能。在自建数据集和公开的 300 头牛只数据集上进行算法验证, 准确率分别达到 88.42% 和 86.67%, 表明了算法的有效性和优越性。将算法移植到嵌入式平台上进行性能测试, 运行速度达到 15 f/s。相比其他算法, 本文算法在牛脸识别上有更好的识别效果。

关键词: 牛脸识别; 特征融合; 动态自适应

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-1298(2026)08-0268-10

OSID:



Cow Face Recognition Algorithm Based on Element-by-element Dynamic Fusion and Adaptive Loss Function

QI Yongsheng^{1,2} ZHAO Jincheng¹ LIU Liqiang^{1,2} LI Yongting^{1,2} WANG Zhaoxia^{1,2}

(1. College of Electric Power, Inner Mongolia University of Technology, Hohhot 010080, China

2. Inner Mongolia Autonomous Region University Smart Energy Technology and Equipment Engineering Research Center, Hohhot 010080, China)

Abstract: Aiming at the problems of poor sensitivity of feature learning, serious interference of complex redundant information in pasture background and difficulty in distinguishing high similarity individuals in cattle recognition task, a cattle face recognition method based on element-by-element dynamic fusion and adaptive loss function was proposed. Firstly, the StarNet architecture was used to redesign the feature extraction network, and an efficient network integrating multi-level features was constructed to extract key visual information and enhance the sensitivity of the model to subtle differences. Secondly, an element-by-element feature fusion module was proposed to screen low-dimensional features and perform weighted fusion on different feature maps to ensure that important features are retained, while irrelevant or redundant features were suppressed and removed. Finally, a dynamic adaptive ArcFace Loss was designed to adaptively adjust the angle boundary of ArcFace Loss, so as to enhance the model's ability to distinguish high similarity samples, and adaptively balance the feature distribution between different categories, which significantly improved the overall recognition performance. The algorithm was verified on the self-built data set and the public data set of 300 cattle. The experimental accuracy reached 88.42% and 86.67%, respectively, indicating the effectiveness and superiority of the algorithm. The

收稿日期: 2024-12-27 修回日期: 2025-02-27

基金项目: 国家自然科学基金项目(62363029)、内蒙古科技计划项目(2020GG0283、2021GG0256)、内蒙古自然科学基金项目(2022MS06018、2024QN06020)、呼和浩特市高校院所协同创新项目(XTCX2023-16)、自治区直属高校基本科研业务费项目(ZTY2024024)和呼和浩特市科技创新领域人才项目(2023RC-联合体-10)

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, E-mail: qys@imut.edu.cn

通信作者: 刘利强(1975—),男,教授,博士,主要从事计算机视觉和新能源发电技术研究, E-mail: llqiang@imut.edu.cn

running speed reached 15 f/s, and it was applied to complex environment testing. Both of them achieved high accuracy, which proved that the algorithm had high stability and robustness. Compared with other algorithms, the proposed algorithm had better recognition effect on cattle face recognition.

Key words: cattle face recognition; feature fusion; dynamic adaptation

0 引言

肉类需求的持续增长推动畜牧业向集约化、专业化转型,牲畜的自动化精准管理与产品质量可追溯^[1]成为行业发展核心需求。而牛只个体的精准识别是实现这一目标的前提^[2-4]。传统身份识别技术如截耳法、耳标法和 RFID 电子耳标法^[5-7]等,会随着成长难以保障有效性和可靠性。因此,亟需探索更加高效的牛只身份识别技术。

随着深度学习的发展,神经网络因其强大的特征提取能力开始在牛只识别领域得到应用。例如,文献[8]研究使用 DenseNet-201、ResNet50 对收集的牛只图像进行交叉验证。XU 等^[9]提出一种将嵌入特征从无效的嵌入空间驱除,从而增强了模型提取身份特征的能力。XU 等^[10]提出的 CattleFaceNet 方法用于识别牛只,实现无损识别。YING 等^[11]提出一种在 ResNet50 主干网络的结构中应用三重注意力机制模块,但会导致使用更多的计算资源和训练时间。实际应用中,识别算法需在嵌入式系统中部署,而嵌入式系统要求基于深度学习的卷积神经网络具有有效的内存、少量的权重表示和低成本的运算量。

此外特征提取过程中,有效特征提取尤为关键。WANG 等^[12]提出一种基于区域注意力网络的识别方法,旨在自适应的捕捉面部不同区域的重要性。ZHAO 等^[13]提出一种局部注意力模块用于引导网络关注局部特征,从而减轻冗余干扰。LIU 等^[14]提出采用 3 种 ResNet50 模型提取面部局部及整体特征,融合映射为 3 个嵌入特征以实现有效提取。付辰伏等^[15]提出一种 Biformer 注意力机制,用在 Backbone、Neck 端之间,以查询自适应的方式关注小面积牛只外观特征标记,聚焦于牛只关键区域。刘世锋等^[16]提出一种自适应阈值背景分离方法进行图像预处理,但需提前对大量数据进行处理,过度依赖参数设置。以上方法均未涉及冗余特征的剔除。

深度学习在面部识别领域应用取得了显著进展,损失函数的优化是关键因素之一。传统损失函数通常基于固定的边距设置,如 DENG 等^[17]提出的 ArcFace 和 WANG 等^[18]提出的 CosFace,可提升类别判别能力,但实际应用有局限性。为了应对这一

挑战,HUANG 等^[19]提出的基于样本难度动态调整学习策略的方案,使模型由易向难逐步推进。此外,ZHANG 等^[20]提出通过调整余弦边距,使得边距随训练进程适配样本分布,提升识别精度。张惠莉等^[21]提出一种自适应权重多任务损失函数联合 KL 散度损失函数与 WingLoss 损失对模型监督。DOHYUNG 等^[22]提出的 Fair Loss 通过自适应加权动态调整样本权重,缓解数据集类别不平衡问题。WANG 等^[23]提出一种融合三元组损失和交叉熵损失以增强网络区分相似个体的能力,并采用 BNNeck 降低损失冲突,提升网络识别能力。现有边距损失函数虽在样本复杂性与类别不平衡处理上取得进展,但在样本自适应性、大规模数据集小样本区分及类别差异敏感性方面仍存不足。

针对以上问题,本文提出一种结合逐元素动态融合模块与自适应损失函数的牛只面部识别算法。基于 StarNet 架构,设计多层次特征融合网络,提炼关键视觉信息;提出逐元素动态特征融合模块,抑制冗余特征;动态调整损失函数,提升高相似度样本的区分能力,增强识别性能;为方便后期在牛场实际环境中部署应用,将算法移植到嵌入式平台 Jetson Xavier NX 上进行测试和实验,以验证算法的有效性和优越性。

1 牛脸识别算法

1.1 算法总体框架

算法总体逻辑如图 1 所示,主要包含 3 个部分:①依据 StarNet 架构对特征提取网络进行重新设计,通过实验构建了[3,3,12,5]最优特征提取网络结构,引入深度可分离卷积-重参数化卷积相结合模块,增强特征学习的敏感性并降低整体模型的复杂性。②在主干网络的浅层,提出一种逐元素动态特征融合模块,通过逐元素比较不同通道的特征图进行筛选,再将子特征图动态特征融合,实现对冗余信息去除,增强关键面部细节的表达能力。③设计一种自适应损失函数,根据样本难易,进行自适应调整损失函数的角度加成,更好地适应类间特征分布的不规则性,解决高相似度区分困难的问题。

1.2 主干特征提取网络

原始主干网络为 MobilenetV1,但原网络存在冗

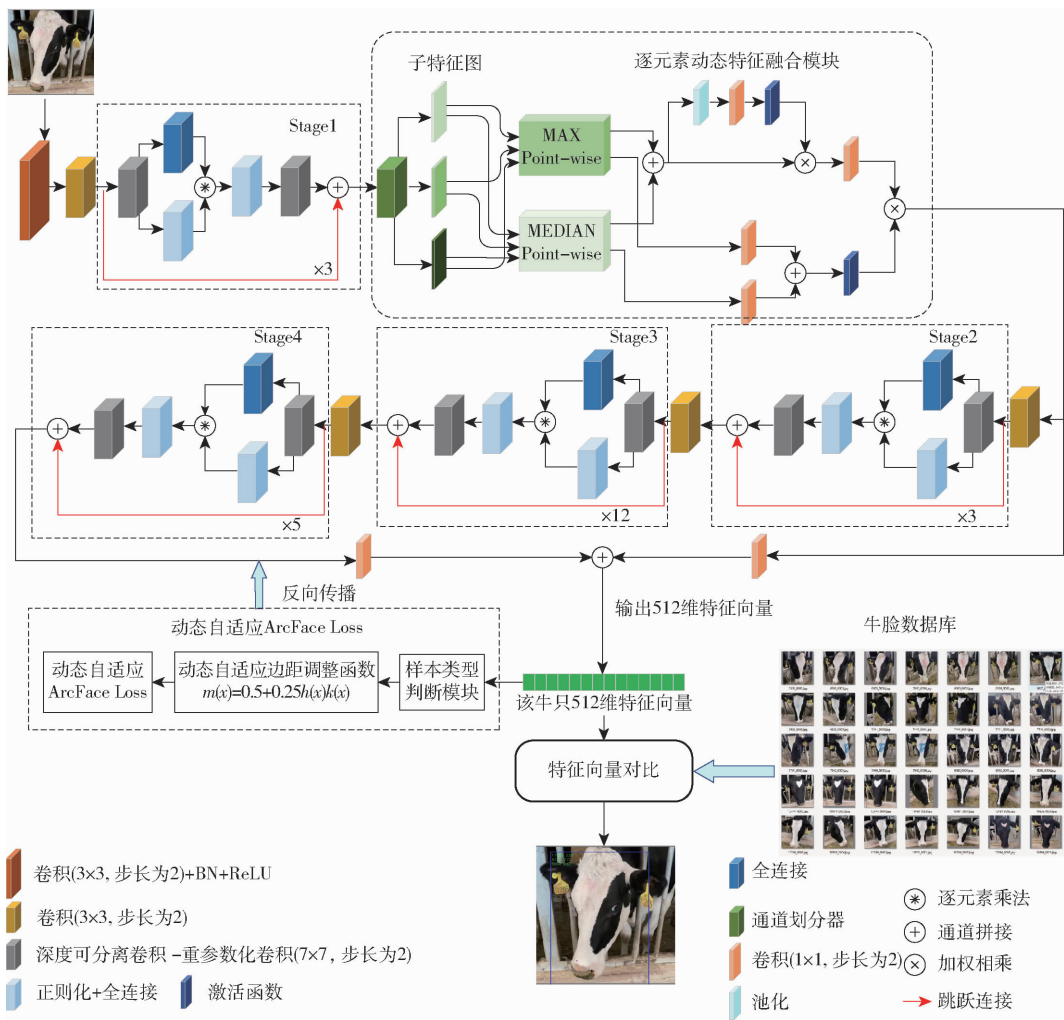


图1 算法结构

Fig.1 Algorithm structure diagram

余计算,且浮点运算量较大,对牛只面部的细微特征敏感性较差,缺乏对全局上下文的整合。在执行牛只识别分类任务中,牛只面部存在许多细微特征,包括纹理、形状、斑点、颜色等,这些特征需要高维特征向量来表示和区分。因此将主干网络设计为一种改进的 StarNet_Cattle。

在传统的 StarNet 中,星操作通过逐元素相乘实现多个特征的非线性融合,并随着网络的深度增加逐渐扩展特征空间。因此通过网络结构对比实验的方式确定最优结构。在牛只识别任务中,需要区分细微差异,比如微小的花纹变化、环境干扰等,这些细节在低维特征中无法有效表示,而深维度的特征可以提取出更细致的信息。因此通过参数量、浮点运算量、准确率、验证率 4 项指标综合考虑,选择最优的结构。最终确定本文最优结构为[3,3,12,5]。

但该结构在捕捉局部特征和增强特征表达能力方面存在局限性,尤其是在捕捉全局上下文信息方面。因此,在每个模块中加入了全局平均池化层(Global average pooling, GAP),通过对每个通道的

空间信息进行平均操作,将整个图像的全局特征提取出来,如图 2 所示。全局平均池化能够有效地捕获图像的全局语义信息,压缩每个通道空间维度,从而为网络提供关于整个图像的全局概览,而不仅仅是局部细节。此外,全局平均池化通过减少特征图维度,避免了传统全连接层带来的计算负担和参数增加,从而简化了网络结构并提升了推理效率。

为进一步提升特征提取网络的性能,引入了深度可分离-重参数化卷积(DW-RepConv)模块,结合了深度可分离卷积(DWConv)和重参数化卷积(RepConv)的优点。在训练阶段,DW-RepConv 采用多分支结构来提取多尺度特征,有效解决了单一深度可分离卷积可能导致特征表达能力下降的问题。具体来说,虽其计算量低,但易削弱模型的特征提取能力;而重参数化卷积通过引入多分支结构,在训练阶段增强了特征提取和权重优化效果,从而提升模型性能。推理阶段,DW-RepConv 会将训练时的多分支结构“折叠”为单分支结构,以显著提高推理效率,同时在保持模型性能的前提下优化计算成

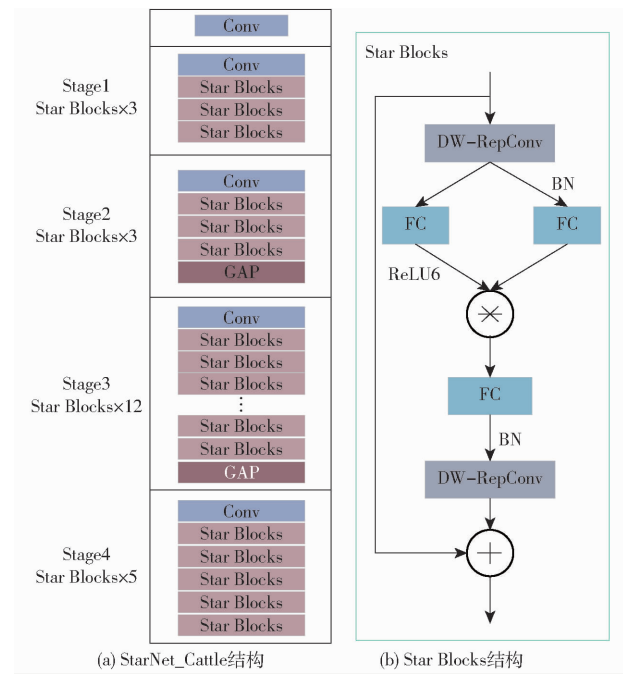


图 2 StarNet_Cattle 网络设计策略

Fig. 2 StarNet_Cattle network design strategy

本。改进的分支结构如图 2b 所示。

1.3 逐元素动态特征融合

从本质上说,提取直接融合的特征常常包含大

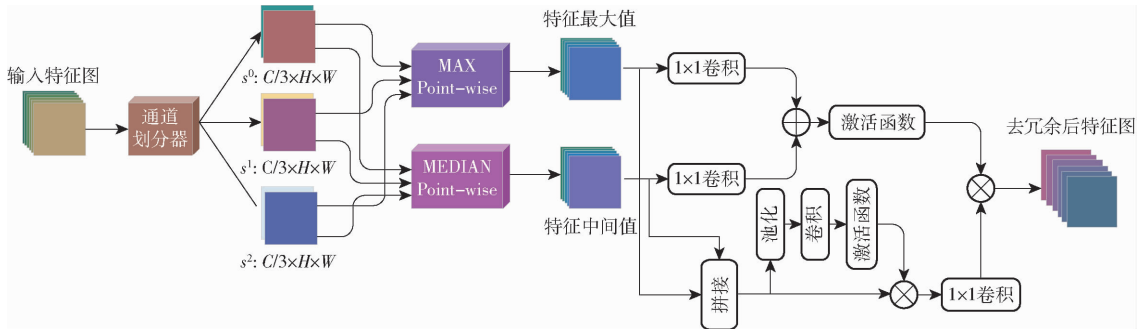


图 3 逐元素动态特征融合模块

Fig. 3 Element-by-element dynamic feature fusion module

得到 2 个子特征图后需对 2 个特征进行融合,但融合过程中,为子特征图选择合适权重是一个关键问题。若权重选择不合适,可能会导致特征偏向某些不重要或无用的特征,而忽略关键特征,导致模型性能下降。因此,提出一种动态特征选择权重的策略。将获得的 2 个子特征图 mdn 和 max 在通道维度上拼接成特征图,为保证后续模块能够直接使用融合特征,设计一种通道升维机制将通道数量扩展到原始数量。传统的扩展机制通过 1×1 卷积进行升维,而本文设计的机制采用全局通道信息 w_{ch} 进行引导升维,该信息通过串联全局平均池化 (AVGPool)、卷积层 (Conv1) 和 Sigmoid 激活函数得到。 w_{ch} 计算式为

$$w_{ch} = \text{Sigmoid}(\text{Conv1}(\text{AVGPool}(\max, \text{mdn}))) \quad (2)$$

量冗余信息,模糊了关键的面部细节。由于包含冗余细节的背景区域,这个问题在现实条件下变得更为凸显。冗余信息的存在会对牛只面部识别任务提出挑战,并可能使系统性能恶化。

为此,提出一种逐元素动态融合模块来消除融合前每个尺度的弱特征,可减少参数量,从而实现更高效、更紧凑的架构。此外,网络的泛化能力通过该方法从相对少量的特征图中提取更多信息特征的能力得到增强。

如图 3 所示,逐元素动态特征融合模块的过程是输入特征图 $f = \{1, 2, \dots, C\} \in \mathbf{R}^{HW}$, 其中 C, H, W 分别表示特征通道数量、高度和宽度。随后通道被分为 3 个不同的子组 (s^0, s^1, s^2) $\in \mathbf{R}^{\frac{C}{3}HW}$ 。对于这 3 个不同的子特征图,通过逐元素比较这 3 个子特征图,得到对应位置中最小值 $\min(x_{chw}^{s^0}, x_{chw}^{s^1}, x_{chw}^{s^2})$, 简称为 $\min$; 中间值 $\text{mdn}(x_{chw}^{s^0}, x_{chw}^{s^1}, x_{chw}^{s^2})$, 简称为 mdn ; 最大值 $\max(x_{chw}^{s^0}, x_{chw}^{s^1}, x_{chw}^{s^2})$, 简称为 $\max$ 。之后,该模块保留最高和中间的特征,并丢弃 3 个子特征图中相同位置的最低特征,从而实现去除冗余信息。其中间特征计算式为

$$\text{mdn} = s^0 + s^1 + s^2 - \max - \min \quad (1)$$

融合后的特征使用全局通道信息 w_{ch} 校准,随后,使用 1 个 1×1 的卷积对特征图进行选择,得到全局空间通道特征 F_c , 计算公式为

$$F_c = \text{Conv1}[w_{ch} \otimes (\max; \text{mdn})] \quad (3)$$

全局通道信息引导卷积层保留了重要特征,抑制了信息量较低的特征,从而实现有效的特征优化。

为建立子特征图之间的空间依赖关系,通过 1×1 卷积后的特征拼接与 Sigmoid 激活得到全局空间信息 w_{sp} , 计算公式为

$$w_{sp} = \text{Sigmoid}(\text{Conv1}(\text{mdn} \oplus \max)) \quad (4)$$

得到的全局空间信息对全局通道特征进行校准,得到去冗后的特征 $\hat{F}$, 具体计算公式为

$$\hat{F} = w_{sp} \otimes F_c \quad (5)$$

1.4 动态自适应损失函数

ArcFace Loss 通过固定的角度边际 m 优化类内

相似性和类间多样性,提升模型类内变化下的鲁棒性。然而,当前损失函数存在以下局限性:随着类间中心距离增大,其对类内特征分布的控制效果减弱,导致类内特征变得松散。同时,损失函数使用固定的角度边际,不直接对类内距离进行惩罚,这在初期可能导致类间中心在特征空间中趋于均匀分布,随后虽能压缩类内特征空间,但这与实际数据中类间中心随机分布的情况并不一致。

ArcFace Loss 采用的加性角边际损失 L_1 , 公式为

$$L_1 = -\frac{1}{n} \sum_{i=1}^n \lg \frac{e^{\text{sacos}(\theta_{y_i} + m_0)}}{e^{\text{sacos}(\theta_{y_i} + m_0)} + \sum_{j=1, j \neq y_i}^N e^{\text{sacos}\theta_j}} \quad (6)$$

式中 x_i ——第 i 个样本深度特征
 y_i ——第 i 头牛只编号
 n ——一次训练迭代中输入的样本数量
 N ——牛只总类别数
 θ_{y_i} ——第 i 头牛只面部特征 x_i 和其真实类别中心之间夹角
 s ——缩放因子,将嵌入特征 $\|x_i\|$ 进行 L_2 归一化后将其缩放到固定值,取 64
 m_0 ——固定的角度加成,在脸部识别中通常取 0.5

但在牛脸识别中,类内和类间的差异分布可能不均匀。当 m_0 取一个较大值时,类内中心与其他类间中心约束过大;当 m_0 取一个较小值时,类内中心远离其他类中心的约束不足。因此,提出一种动态自适应角度损失函数。

首先根据输入特征的角度判断是否需要调整 m_0 , 为此需引入动态样本评估函数 $h(x)$, 即

$$h(x_i) = \begin{cases} -1 & (\theta_{y_i} + m_0 \leq \min(\text{inter})) \\ 1 & (\theta_{y_i} + m_0 > \min(\text{inter})) \end{cases} \quad (7)$$

式中 $\min(\text{inter})$ ——样本与非真实类别中心之间的最小角度

若 $(x_i + m_0)$ 小于等于最小非真实类别的夹角,称其为易样本,说明样本 x_i 更接近真实类别中心, m_0 对样本影响作用较小,需降低 m_0 的值。若 $(x_i + m_0)$ 大于最小非真实样本的夹角,称其为难样本,说明样本 x_i 在特征空间中更接近某个非真实类别中心,模型更易将其错误分类为该非真实类别,此时 m_0 的值对样本分类影响较大,就需要对 m_0 进行调整,使其更接近类别中心,以达到增强分类效果。

因此对难样本进行动态自适应调整边距,让其更接近自身类别中心。选择 θ_{y_i} 作为动态调整的指标, m 随着 θ_{y_i} 变化而变化。

在深度学习模型中,特征向量与类别中心向量之间的角度 θ_{y_i} 通常通过余弦相似度计算得到,计算式为

$$\cos\theta_{y_i} = \frac{W_{y_i} x_{y_i}}{\|W_{y_i}\| \|x_{y_i}\|} \quad (8)$$

式中 W_{y_i} ——第 i 头牛只第 j 个类别权重向量

在训练过程,需将特征归一化到单位长度,所以 $\cos\theta_{y_i} \in [0, 1]$, 因此 $\theta_{y_i} \in [0, \pi/2]$ 。设计了函数 $k(x)$, 根据 θ_{y_i} 值判断样本是否容易分类,然后进行动态调整边距,公式为

$$k(x) = 1 - \cos\theta_{y_i} \quad (9)$$

其中,函数 $k(x)$ 为单调递增的凸函数。当 θ_{y_i} 值接近 0 时, $k(x)$ 增长缓慢,当 θ_{y_i} 值接近 $\pi/2$ 时, $k(x)$ 增长速度迅速,表明,难样本会被分配到更大的边距。

最终定义一种动态自适应 m 的函数,根据样本类型由 θ_{y_i} 决定 m , 具体公式为

$$m(x) = m_0 + t h(x) k(x) \quad (10)$$

其中, t 为比例因子,用于将 $m(x)$ 调整到适合的范围内。因此将 t 设置为 0.25, 避免因边距过大而导致模型性能下降或无法收敛。

因此,自适应损失函数 L_2 公式为

$$L_2 = -\frac{1}{n} \sum_{i=1}^n \lg \frac{e^{\text{sacos}(\theta_{y_i} + m(\theta_{y_i}))}}{e^{\text{sacos}(\theta_{y_i} + m(\theta_{y_i}))} + \sum_{j=1, j \neq y_i}^N e^{\text{sacos}\theta_j}} \quad (11)$$

2 实验与结果分析

2.1 验证数据集

数据采集于 2023 年 3—12 月呼和浩特市多个牧场,采集时间段为 07:00—09:00 和 17:00—19:00,2 个时段恰逢牛只喂料时间,可降低人为因素对牛只的干扰。

所用采集设备如图 4 所示。对同一牛只左、正、右三脸(3 个角度)同时进行拍摄。通过从不同角度(左脸、正面和右脸)同时捕捉牛只的面部图像,可以全面获取牛只的面部特征,避免单一视角可能带来的识别误差。三角度同时拍摄不仅能捕捉到更多细节,还能提高识别的准确性和鲁棒性,减少因姿势变化或遮挡导致的后期识别失败。采用 3 台海康威视 DS-2CD3626FWD-LPTZ 型高清图像摄像头,该摄像头最大分辨率为 2 560 像素 $\times$ 1 920 像素,可以清晰捕捉牛只面部细节。该相机内置智能红外技术和自动调整红外强度,能在低光、完全黑暗、过度曝光的环境中提供清晰图像。该相机具备较强的防水防尘能力,能在湿润、尘土较多的环境中保证持续的工作稳定性,确保相机长时间稳定运行。

使用采集到的 900 多头肉牛牛脸数据进行模型训练,其中每头牛只包含左、右、正 3 个角度各 3 幅,



图 4 牛只面部采集设备

Fig.4 Cattle face acquisition equipment

总计 5 580 幅图像。为增加样本多样性,通过翻转、放缩、色域变换、改变对比度、增加噪声等数据增强方法对数据集进行扩充,扩充后图像数量为 55 800 幅。按照比例 8: 1: 1 划分为训练集、测试集、验证集,具体划分数量如表 1 所示。为实现开放集合牛只面部的识别,在测试集和验证集中引入未训练的牛只样本。

表 1 牛只面部数据集划分

Tab.1 Division of cattle face dataset

数据集种类	训练集	验证集	闭集测试集	开集测试集
数据集数量	44 640	5 580	5 580	7 254

2.2 实验环境及算法参数设置

为满足后期算法可移植性的要求,首先在服务器进行训练。操作系统为 Ubuntu18.04 Linux,硬件环境配置为:i7 - 11700@ 2.50 GHz CPU、NVIDIA GeForce RTX 3090 显卡、运行内存 32 GB。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;Python 3.6。模型训练策略:使用 Adam 优化器进行优化,总迭代次数为 100,迭代动量系数为 0.9,批量大小设置为 64,权重衰减系数为 5×10^{-4} ,初始学习率为 0.01,学习下降率采用余弦退火的动态调整的方式。

2.3 评价指标

使用识别准确率(Accuracy)、验证率(Validation rate)、帧率、参数量和模型内存占用量作为模型性能评价指标。

2.4 训练结果

模型在自制牛脸数据集性能训练结果如图 5 所示。在训练过程,使用角度作为度量方式,即将特征向量与动态类中心向量间的夹角作为度量标准,并将验证集损失最小值作为最优模型权重。在训练过程,随着训练轮数的不断增加,损失值逐渐降低,最终趋于稳定。在模型前 20 轮,训练损失和验证损失均快速下降,说明模型在前期学习了有效的特征,第 20 轮后,损失值趋于平稳,模型趋于收敛,性能达到最优。之后,验证与训练损失的趋势基本一致,表明

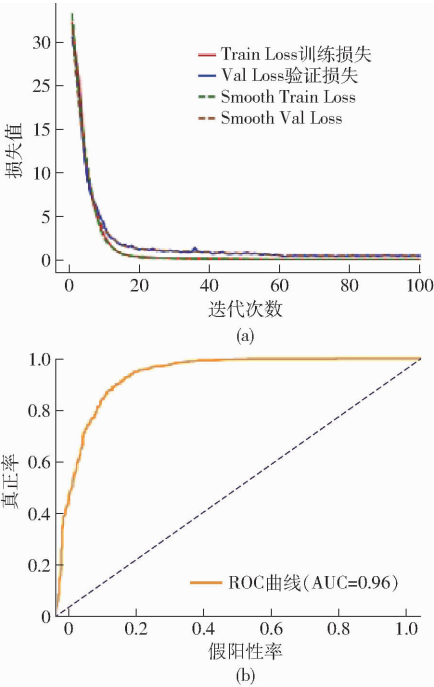


图 5 模型训练曲线

Fig.5 Model training curves

模型没有严重的过拟合现象。

对于每个确定的阈值,高于阈值的实例被认为是正类,低于阈值则认为是负类。为描述这一变化引入 ROC 曲线,可以用于评价一个分类模型。ROC 曲线是以假阳性率(FPR)为横轴,真正率(TPR)为纵轴,针对每一个确定的阈值计算相应的 FPR 和 TPR 数值点,将所有点相连接就能够得到相应的 ROC 曲线,可以很好地反映模型性能,当曲线越靠近左上角时,说明模型的灵敏度越高,性能越好。测试结果如图 5b 所示。

2.5 消融实验

为验证本文算法改进的有效性,在自制牛脸数据集上进行消融实验,并在原主干网络中的 MobileNetV1 和替换的主干网络 StarNet_Cattle 进行消融实验,去除逐元素动态融合模块与自适应损失函数比较算法准确率。消融实验结果如表 2 所示。

由表 2 可知,主干网络改进为 StarNet_Cattle 后,参数量有明显降低,网络识别的准确率也有提升,证明将数据映射到高维空间,数据的不同类别往往可以变得更加线性可分,可以进一步提高准确率。在网络中增加了逐元素动态特征融合模块后,将特征进一步筛选,有效地选择和增强关键信息,突出最为重要的低维度特征,从而减少冗余信息的干扰,准确率提升 2.11 个百分点,验证率提升 2.07 个百分点。此外,自适应损失函数的加入,增强了样本类内紧凑性和类间可分离性,准确率提升 2.03 个百分

表 2 消融实验结果

Tab.2 Ablation experiment results

主干网络	参数量	准确率/%	验证率/%	帧率/(f·s ⁻¹)
MobilenetV1	1.658 × 10 ⁷	84.05	79.72	35.87
MobilenetV1 + 逐元素动态融合模块	1.745 × 10 ⁷	84.93	81.24	35.16
MobilenetV1 + 逐元素动态融合模块 + 自适应损失	1.835 × 10 ⁷	85.92	82.36	34.58
StarNet_Cattle	8.725 × 10 ⁶	84.28	82.28	36.75
StarNet_Cattle + 逐元素动态融合模块	9.325 × 10 ⁶	86.39	84.35	36.02
StarNet_Cattle + 逐元素动态融合模块 + 自适应损失	9.467 × 10 ⁶	88.42	87.68	35.95

点,验证率提升 3.33 个百分点。3 种改进措施的加入均提升了模型准确率,相比原网络有明显提升,验证了本文算法的有效性和先进性。

2.6 不同算法性能测试

2.6.1 不同算法性能测试

为进一步评估模型性能,将本文算法与当前典型的识别算法进行对比,包括: FaceNet^[24]、MobileFaceNet^[25]、ArcFace^[17]、SK _ ResNet^[26]、VanillaFaceNet^[27]、PocketNetS^[28]。其中,FaceNet 为应用最为广泛的面部识别算法,主干特征提取网络使用 Inception - ResNet; MobileFaceNet 为改进的轻

量化面部识别网络;SK_ResNet 为 ResNet - 50 的改进算法,使用不同数量的 SK_Bottleneck 融合多个感知信息,并在多尺度上提取牛只面部特征; VanillaFaceNet 主干特征网络使用 Vanillanet 作为主干特征提取网络,融合通道和位置信息的高效的注意力机制,降低干扰信息。PocketNetS 采用对传统卷积神经网络进行精简,去除不必要的卷积层和过多的特征图实现在较小存储开销下保持较好的精度。各算法均与本文算法的训练环境、训练轮次相同。得到各算法的准确率、模型内存占有量和帧率如表 3 所示。

表 3 不同算法测试结果

Tab.3 Test results of different algorithms

算法	参数量	自建数据集			公开 300 头牛只数据集		
		模型内存占有量/MB	准确率/%	帧率/(f·s ⁻¹)	模型内存占有量/MB	准确率/%	帧率/(f·s ⁻¹)
FaceNet ^[24]	2.279 × 10 ⁷	87.40	85.51	23.56	86.38	84.39	23.38
MobileFaceNet ^[25]	1.003 × 10 ⁶	4.80	78.63	38.25	4.61	75.21	38.17
ArcFace ^[17]	1.658 × 10 ⁷	67.40	83.72	31.28	66.80	80.98	31.09
SK_ResNet ^[26]	4.359 × 10 ⁷	22.56	84.22	24.86	21.68	81.54	24.65
VanillaFaceNet ^[27]	1.051 × 10 ⁸	101.70	88.21	24.72	93.20	85.32	24.32
PocketNetS ^[28]	9.600 × 10 ⁵	4.11	80.74	37.61	3.83	75.53	36.05
本文算法	9.467 × 10 ⁶	90.30	88.42	34.74	37.63	86.67	34.19

由表 3 可知,本文算法识别精度明显高于其他算法,且参数量小于大部分算法,算法性能优良。算法精度方面,逐元素动态特征融合模块与自适应损失函数可以有效地提升牛脸识别精度。MobileFaceNet 模型内存占用量和参数量相对较小,但是准确率相对较低。相比较为复杂的 Arcface 模型,本文算法能够保证网络轻量化的同时,拥有较快的识别速度。SK _ ResNet 由于使用多个 SK _ Bottleneck 来融合多个感知信息,参数量以及模型内存占用量都高于本文算法,但识别准确率与识别速度稍低于本文算法。VanillaFaceNet 参数量和模型内存占用量都相对较高,但其准确率略低于本算法。PocketNetS 作为轻量化模型,但因其对网络精简在牛只识别中无法提取大量有效特征造成准确率较低。综上,本文算法在识别准确率上具有明显优势。

由表 3 可知,基于自建数据集比基于公开数据

集识别效果的准确率高。因自建数据集数据量丰富,且牛只均为低头状态可提取到较为完整的特征,而公开数据集牛只均为昂头状态,存在提取特征不全面的不足。且本文算法在各数据集上均得到较高的准确率,证明了算法的有效性。

2.6.2 主干特征提取网络对比测试

在设计和选择 StarNet 结构时,需要综合考虑任务需求、硬件资源、以及模型特征提取能力。不同的 StarNet 结构在嵌入维度和网络深度上存在差异,适用于不同应用场景。因此需从多方面对比各个结构的优劣。

涉及精细的牛脸特征提取,例如个体身份精准识别,需选择具有更深层次特征提取能力的模型,所以要选择深度较深的网络。对比参数量、浮点运算量、准确率、验证率,选择最合适网络结构。结果如表 4 所示。

表 4 主干特征提取网络结构测试结果

Tab.4 Test results of backbone feature extraction network structure

变形	图像嵌入	深度	参数量	浮点运算量	准确率/%	验证率/%
S0	16	[1,1,3,1]	1.552×10^6	5.179×10^{10}	78.40	77.82
S1	24	[2,2,8,3]	4.257×10^6	2.241×10^{11}	79.28	78.25
S2	32	[1,2,6,2]	5.521×10^6	2.932×10^{11}	79.75	79.01
S3	32	[2,2,8,4]	7.596×10^6	4.111×10^{11}	81.40	79.86
S4	32	[3,3,12,5]	8.743×10^6	5.660×10^{11}	83.82	82.28
S5	64	[6,6,14,8]	1.027×10^7	6.590×10^{11}	83.85	82.32
S6	32	[3,3,12,5,5]	1.224×10^7	9.350×10^{11}	84.02	82.56

由表 4 可知,随网络深度的逐步增加,模型参数量、浮点运算量、准确率和验证率均呈上升趋势。然而,当网络结构深度为[6,6,14,8]时,尽管参数量和浮点运算量显著增加,但准确率相较于前一个深度的提升幅度有限,仅从 83.82% 增长到 83.85%,但参数量从 8.743×10^6 增加到 1.027×10^7 。继续加深网络准确率提升不明显,无法满足深度和性能的平衡。因此,最终选择深度[3,3,12,5]作为当前任务最优网络结构,从而满足在复杂度和准确性之间较优的平衡。

2.6.3 逐元素动态融合模块有效验证实验

为验证提出的逐元素动态融合特征模块的有效性,图 6 可视化了原始主干网络和增加模块后的主干网络特征图,图中颜色柱表示将特征值的强度映射到 0~255 颜色范围,0 表示颜色柱底部,表示特征值最小,255 表示特征值最大。与原始特征网络相比,增加模块后明显有颜色差距,浅色表示冗余特征被抑制,红色表示重要特征被增强,且颜色深度表示特征有用性。可视化结果证明了使用该模块的有效性。

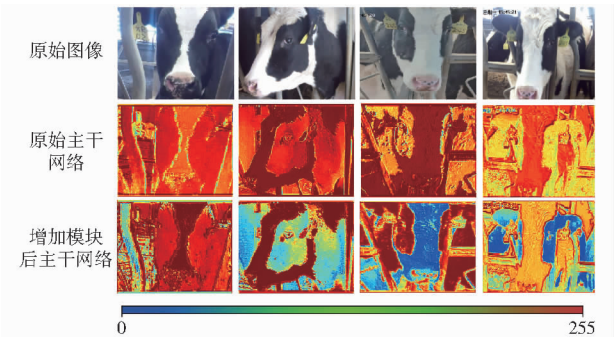


图 6 中间层特征可视化结果

Fig.6 Visualization results of middle layer features

2.6.4 损失函数测试

为证明所采用的动态损失函数的优越性,将原模型结合改进的动态自适应 ArcFace Loss 与其他当前的损失函数组合模型进行对比分析,严格控制测试参数,准确率结果如表 5 所示。

表 5 不同损失函数测试结果

Tab.5 Test results of different loss functions %

损失函数	训练集准确率	测试集准确率
三元组损失	91.47	83.32
AM Softmax	92.11	83.47
ArcFace	93.12	84.26
自适应损失	96.08	85.32

由表 5 可知,使用 AM Softmax、ArcFace、自适应 ArcFace 改进网络,不仅在训练时准确率较三元组损失提升,测试时准确率也有所提升。说明使用角度间隔作为度量更为优秀,并增强模型效果。使用 ArcFace 损失函数改进网络模型,证明角度空间中直接进行最大化决策边界能够使模型具有一定判别能力,能够获得更有效的特征,准确率较 AM Softmax 有所提升,在测试集上提升 0.79 个百分点,达到 84.26%。而本文使用的动态自适应损失相比于 ArcFace Loss,准确率提升 1.06 个百分点,达到 85.32%,说明在决策边界上判断力更优,从而使得特征区分更明显。

2.7 特征可视化

为更好地理解本文识别效果,使用 Grad-CAM (梯度加权类激活映射)可视化算法对图像注意力区域,输入图像为编号 0005、6327、6606、6631 的 4 幅相似牛脸图像。

Grad-CAM 通过分析梯度的信息捕获模型对图像中不同牛脸图像部分的关注程度,从而生成一个良好的视觉解释,指示模型在做出预测时对图像的哪些部分有更强的关注,可以帮助理解模型在牛脸识别任务中为何做出特定的预测。Grad-CAM 可视化结果如图 7 所示,红色表示该区域中存在高激活,蓝色表示预测类别弱激活。由图 7 可以发现,当使用原算法时,模型较少关注牛只面部区域。在进行优化操作后,本文算法活跃区域更大,对牛只面部关注度更高。

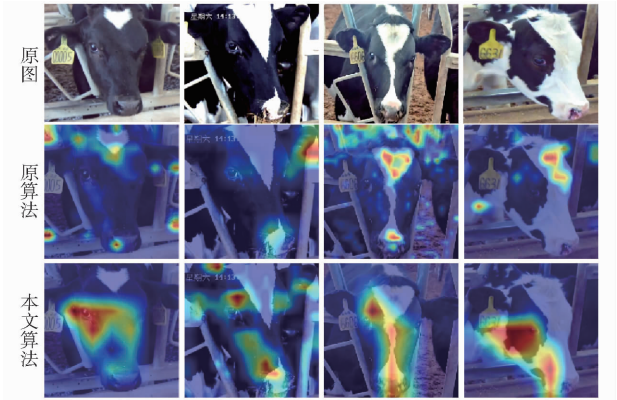


图 7 不同网络牛只面部可视化

Fig. 7 Visualization of cattle face in different networks

3 算法移植

为保证算法在提升检测精度的同时满足移植要求,将模型移植到 Jetson Xavier NX 嵌入式平台进行测试。使用 RetinaFace 检测算法进行牛脸检测,将检测的区域进行裁剪,确定需要识别的牛只。之后将其输入到后端算法进行牛只识别,测试平台组成如图 8 所示。



图 8 嵌入式平台实物图

Fig. 8 Embedded platform to build physical map

如图 9 所示,在检测框左上角有 3 组数字,表示识别该牛只具体编号以及准确率。经过大量运行测试,测试结果表明,本算法在嵌入式系统上运行速度为 15 f/s,验证本文算法实时性。

选取强光、弱光、相机抖动导致的模糊场景进行识别,从而体现算法在不同场景下应用的真实性,识别结果如图 10 所示。图 10 中横向图像为 1 组,每幅图像左上角标号为牛只对应身份编号。由 3 组图像左上角识别结果与其耳标编号相验证,可知最终牛脸



图 9 嵌入式平台下牛只识别结果

Fig. 9 Cattle recognition results under embedded platform



图 10 不同干扰场景下牛只识别结果

Fig. 10 Cattle recognition results under different interference scenarios

识别结果均正确,验证了本文算法稳定性和鲁棒性。

4 结束语

提出一种基于逐元素动态融合与自适应损失函数相结合的牛只面部识别方法,解决了传统牛脸识别任务中存在特征学习敏感性差、牧场背景复杂冗余信息干扰严重及高相似度个体区分困难的问题。所提算法在实时牛脸识别任务中,可保证准确度和运行速度实时性要求。

参 考 文 献

[1] BAO J, XIE Q. Artificial intelligence in animal farming; a systematic literature review[J]. Journal of Cleaner Production, 2022, 331: 129956.

[2] SUN Y, HUO P, WANG Y, et al. Automatic monitoring system for individual dairy cows based on a deep learning framework that provides identification via body parts and estimation of body condition score[J]. Journal of Dairy Science, 2019, 102(11): 10140 - 10151.

[3] LI X, WANG M. Study on the livelihood capital level, structural characteristics, and coupling coordination degree of Chinese beef cattle farmers[J]. Agriculture, 2023, 13(7): 1371.

- [4] BERGMAN N, YITZHAKY Y, HALACHMI I. Biometric identification of dairy cows via real-time facial recognition[J]. *Animal*, 2024, 18(3): 101079.
- [5] AWAD A I. From classical methods to animal biometrics: a review on cattle identification and tracking[J]. *Computers and Electronics in Agriculture*, 2016, 123: 423 – 435.
- [6] ZIN T T, PWINT M Z, SEINT P T, et al. Automatic cow location tracking system using ear tag visual analysis[J]. *Sensors*, 2020, 20(12): 3564.
- [7] ZHANG T, CAO C, YU H, et al. Design and implementation of dairy food tracking system based on RFID[C] // 2020 International Wireless Communications and Mobile Computing (IWCMC), 2020: 2199 – 2203.
- [8] DE LIMA WEBER F, DE MORAES WEBER V A, MENEZES G V, et al. Recognition of Pantaneira cattle breed using computer vision and convolutional neural networks[J]. *Computers and Electronics in Agriculture*, 2020, 175: 105548.
- [9] XU X, DENG H, WANG Y, et al. Boosting cattle face recognition under uncontrolled scenes by embedding enhancement and optimization[J]. *Applied Soft Computing*, 2024, 164: 111951.
- [10] XU B, WANG W, GUO L, et al. CattleFaceNet: a cattle face identification approach based on RetinaFace and ArcFace loss [J]. *Computers and Electronics in Agriculture*, 2022, 193: 106675.
- [11] YING Xiaoyi, ZHAO Jizheng, YANG Lingling, et al. Re-identifying beef cattle using improved AlignedReID + [J]. *Transactions of the CSAE*, 2024, 40(18): 132 – 146.
- [12] WANG K, PENG X, YANG J, et al. Region attention networks for pose and occlusion robust facial expression recognition[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4057 – 4069.
- [13] ZHAO Z, LIU Q, WANG S. Learning deep global multi-scale and local attention features for facial expression recognition in the wild[J]. *IEEE Transactions on Image Processing*, 2021, 30: 6544 – 6556.
- [14] LIU Y, DAI W, FANG F, et al. Dynamic multi-channel metric network for joint pose-aware and identity-invariant facial expression recognition[J]. *Information Sciences*, 2021, 578: 195 – 213.
- [15] 付辰伏, 任力生, 王芳. 基于改进 YOLO v8 的牛只行为识别与跟踪方法[J]. *农业机械学报*, 2024, 55(5): 290 – 301.
FU Chenfu, REN Lisheng, WANG Fang. Method for cattle behavior recognition and tracking based on improved YOLO v8[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(5): 290 – 301. (in Chinese)
- [16] 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)
- [17] DENG J, GUO J, XUE N, et al. Arcface: additive angular margin loss for deep face recognition[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 4690 – 4699.
- [18] WANG H, WANG Y, ZHOU Z, et al. Cosface: large margin cosine loss for deep face recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 5265 – 5274.
- [19] HUANG Y, WANG Y, TAI Y, et al. Curricularface: adaptive curriculum learning loss for deep face recognition[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 5901 – 5910.
- [20] ZHANG X, ZHAO R, QIAO Y, et al. Adacos: adaptively scaling cosine logits for effectively learning deep face representations[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 10823 – 10832.
- [21] 张惠莉, 王光远, 员玉良, 等. 基于 FCM-SimCC 的猪只面部关键点定位方法[J]. *农业机械学报*, 2025, 56(4): 397 – 407.
ZHANG Huili, WANG Guangyuan, YUAN Yuliang, et al. Pig facial landmark detection method based on FCM-SimCC[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(4): 397 – 407. (in Chinese)
- [22] DOHYUNG K, SUNGHO P, SUNHEE H, et al. Fair classification by loss balancing via fairness-aware batch sampling[J]. *Neurocomputing*, 2023, 518: 231 – 241.
- [23] 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.
- [24] 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.
- [25] LU Z, ZHOU C, WANG H. A face detection and recognition method built on the improved MobileFaceNet[J]. *International Journal of Sensor Networks*, 2024, 45(3): 166 – 176.
- [26] GONG H, PAN H, CHEN L, et al. Facial recognition of cattle based on SK-ResNet[J]. *Scientific Programming*, 2022, 2022: 1 – 10.
- [27] 栾浩天, 齐咏生, 刘利强, 等. VanillaFaceNet: 一种高精度快速推理的牛脸识别方法[J]. *农业工程学报*, 2024, 40(18): 120 – 131.
LUAN Haotian, QI Yongsheng, LIU Liqiang, et al. VanillaFaceNet: a high-precision and rapid inference for bovine face recognition[J]. *Transactions of the CASE*, 2024, 40(18): 120 – 131. (in Chinese)
- [28] BOUTROS F, SIEBKE P, KLEMT M, et al. Pocketnet: extreme lightweight face recognition network using neural architecture search and multistep knowledge distillation[J]. *IEEE Access*, 2022, 10: 46823 – 46833.