

刘峰, 吴文杰, 刘小磊, 等. 计算机视觉与深度学习在猪只识别中的研究进展[J]. 华中农业大学学报, 2023, 42(3): 47-56.
DOI: 10.13300/j.cnki.hnlkxb.2023.03.006

计算机视觉与深度学习在猪只识别中的研究进展

刘峰^{1,2,3}, 吴文杰¹, 刘小磊^{2,3,4}, 王欣然¹,
方亚平¹, 李国亮^{1,2,3}, 杜小勇^{1,2,3,4}

1. 华中农业大学信息学院/农业农村部智慧养殖技术重点实验室/
农业智能技术教育部工程研究中心/湖北省农业大数据工程技术研究中心, 武汉 430070;

2. 华中农业大学深圳营养与健康研究院, 深圳 518000;

3. 中国农业科学院深圳农业基因组研究所/岭南现代农业科学与技术广东省实验室深圳分中心, 深圳 518000;

4. 农业动物遗传育种与繁殖教育部重点实验室, 武汉 430070

摘要 探索人工智能领域新技术与生猪养殖相结合, 是当前智慧养殖领域的一个重要研究方向。其中, 如何自动地识别猪只个体身份与行为, 是当前生猪养殖行业要解决的一个关键问题。为推动计算机视觉和深度学习技术在猪只健康状态智能化监测方面的应用, 本文先分析了基于计算机视觉与深度神经网络的人的身份及行为识别模型的研究进展, 然后对利用计算机视觉与深度神经网络识别猪只个体身份及行为的方法进行了归纳总结, 并指出已有方法中存在的问题, 最后提出了下一步的重点研究方向: (1) 在猪只运动不可控及关键特征部位受到污染的情况下, 准确提取其身份及行为特征的方法研究; (2) 针对猪只身份及行为特征的基于计算机视觉的原创性深度学习模型的研究; (3) 能够同时检测猪只身份及行为的多任务神经网络的研究; (4) 适用于多场景的基于基础姿态及动作的通用型猪只行为识别方法的研究; (5) 基于边缘计算的猪只个体身份及行为识别的部署方法研究。

关键词 身份识别; 行为识别; 深度学习; 计算机视觉; 智慧养殖; 猪脸识别; 自动监测

中图分类号 S817.3 **文献标识码** A **文章编号** 1000-2421(2023)03-0047-10

当前畜牧业发展正处于转型的关键时期, 生猪产业逐渐由家庭生产向集约型大规模养殖转变。随着传感器技术、视频监控技术、信息通信技术、大数据和人工智能技术的发展, 养猪业逐渐进入智能化养殖时代^[1]。将传统猪场升级为智慧猪场, 实现猪场的精细饲养和疾病防控, 同时完善生猪养殖产业链, 实现从农场到餐桌的数据全流程溯源, 可确保猪肉产品质量安全, 也可预防区域性的动物疫情发生^[2]。实现智能化和精准化养殖的关键在于对猪只个体的身份识别, 因为猪只个体身份识别不仅是健康生产管理的基础, 也是猪肉产品溯源的重要组成部分^[3]。

传统上, 判断猪只健康状况主要根据饲养员的经验, 方法耗时且费力, 随着规模化猪场发展, 需要更客观的智能养猪管理工具以快速判断^[1]。通过行

为识别与分析, 检测与猪只疫病相关的早期异常行为, 可有效防御疫病的大规模传播。由于生猪养殖通常采用群体养殖的方式, 不仅需要识别到猪只的行为, 同时也要检测到猪只个体身份, 真正实现智能化与精准化。因此, 猪只个体身份识别是猪只行为识别的前提^[4]。在动物个体身份精准识别基础上的行为识别问题, 是智能化畜牧养殖领域最重要的基础工作之一。该问题的解决对提升疫情预警效率、降低养殖成本、提高动物福利、提升肉类品质等具有重要的理论与实际意义。

猪只个体身份与行为识别方法, 是利用在计算机视觉及深度学习领域中研究成熟的人的身份、行为检测及识别模型, 结合生猪饲养的自身特征加以改进, 或直接采用猪只图像及视频数据重新训练, 以

收稿日期: 2022-10-31

基金项目: 华中农业大学-中国农业科学院深圳农业基因组研究所合作基金(SZYJY2021011; SYZJY2022001); 国家重点研发计划青年科学家项目(2021YFD1300800)

刘峰, E-mail: liufeng@mail.hzau.edu.cn

通信作者: 杜小勇, E-mail: duxiaoyong@mail.hzau.edu.cn

满足猪只个体身份与行为识别的需要^[5]。近年来,深度学习在各种计算机视觉问题上取得了长足的进步,如对象检测^[6-7]、运动跟踪^[8]、动作识别^[9]、人体姿势估计^[10]和语义分割^[11]。深度学习允许多个处理层的计算模型学习和处理具有多个抽象级别的数据,模仿大脑感知和理解多模态信息,隐性地捕获大规模数据的复杂结构^[12]。大型、高质量、公开可用的标记数据集,并行GPU的深度模型高速训练设备,以及如TensorFlow、theano^[13]、mxnet、pytorch等强大框架的出现,促成深度学习的学习效率明显提升。因此,本文在基于计算机视觉的人的个体身份及行为识别深度模型的研究基础上,归纳猪只个体身份与行为识别的研究进展,并展望该领域未来的研究方向,以期智能化、精准化生猪养殖,尤其是机器视觉和深度学习在养猪行业的应用提供技术支持。

1 个体身份与行为识别中的计算机视觉及深度学习模型

计算机视觉技术最早是根据人类视觉系统、相

机、投影及摄影测量法等基本原理,进行计算机视觉图像的处理^[14]。近几十年来,计算机视觉技术在模式识别、机器学习与深度学习、计算机图形学、3D重构、虚拟现实与增强现实等领域迅速发展,已广泛应用于对象检测、机动车自动导航与驾驶、车牌识别、指纹识别、人脸识别、姿态及行为识别等具体场景^[15-17]。

下面对这些技术在人的个体身份和行为识别方面的一些典型研究工作进行简要介绍。

1.1 基于计算机视觉的个体身份识别深度学习模型

当前针对个体身份识别的深度学习研究,主要是基于人脸进行人的身份的识别。其基本思路:首先检测目标,在一张图像里找到可能存在的目标对象,其次通过深度神经网络提取人脸或身体图片中的特征向量,再次根据特征向量间的欧式距离、球面距离等标准,计算特征向量间的相似度,最后根据相似度来判定这2张图片是否属于同一个体。具体如图1所示。

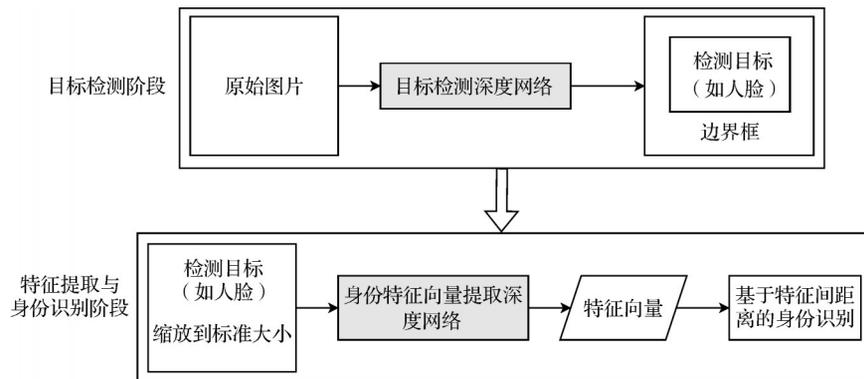


图1 基于深度学习的个体身份识别方法

Fig.1 Identity recognition methods based on computer vision and deep learning

目标检测是身份及行为特征提取的前提与基础。其目的为获得能够覆盖目标对象的矩形边界框的4点坐标,并采用分类的方式得到目标类别。常用的模型与方法有SSD^[18]及其在MobileNet^[19]移动平台上的版本 MobileNet-SSD、YOLO^[20-22]、MTCNN^[23]、R-CNN^[24]等。R-CNN及其改进版本Fast R-CNN^[25]采用了区块特征提取思路,在准确度上相比于YOLO等方法要高,但是需要更多的计算资源,计算速度较慢。在Faster R-CNN^[26]基础上设计的Mask R-CNN模型^[27]具有图像分割功能,这是一般的目标检测模型所不具备的,对于精准目标分割以生成精确特征码应用方面具有重要意义。针对

以Mask R-CNN模型为代表的“先检测再分割”的实例分割方式的时间复杂度高的问题,Wang等^[28]借鉴YOLO中图像网格的思想,提出实例分类的方法,将图像实例分割转换为分类问题,并采用点到点的方式构建深度模型,提升了实例分割的效率与准确度。

在检测到人脸的基础上,再通过特征向量提取及比对来进行人脸识别。2014年,Taigman等^[29]提出了DeepFace模型,采用9层深度神经网络,在超过4 000人共计400万张图片集上进行训练,以提取人脸特征,形成人脸的特征向量,并通过计算不同人脸特征向量之间的距离来衡量相似性。2015年,谷

歌公司提出了人脸识别的FaceNet方法^[30],采用Inception深度网络模型及Triplet损失函数,得到人脸的特征向量,并通过计算人脸特征向量间的欧式距离来判定是否属于同一个人。我国旷视科技在2015年提出了Face++模型^[31],构建了10层的神经网络,分别对人脸图片的4个区域生成特征码,这4个特征码共同用于构建人脸特征向量,根据欧式距离进行相似度计算。2015年百度提出了人脸识别模型^[32],针对不同脸部部位分别生成各个部位的特征向量,然后连接成最终的全局特征向量,并用欧式距离计算脸部相似度。此外,Liu等^[33]针对人脸相似度计算提出了球面距离计算方法。理论上这些生物特征识别技术大部分都可以应用于在猪只个体身份与行为识别领域。但是与人脸识别还有明显不同,由

于猪只识别发生在农场场景,其环境更不可控,如阳光、灰尘以及猪只不受指令的活动等要素;此外,猪比人类衰老得快很多,在猪只个体的快速生长期可能无法持续有效地使用同一个面部识别模型。

1.2 基于计算机视觉的姿态及行为识别深度学习模型

对于人体姿态的检测通常有自上而下与自下而上2种方法。自上而下的识别方式是首先检测出图片中的每个人体对象,再分别检测个体上的每个关键部位点,如图2所示。而自下而上的方式则是首先检测出1张图片上所有的关键部位点,再将这点聚合成人体的不同个体,如图3所示。通常情况下,自上而下的方法具有更高的识别准确率,而自下而上的方法具有更快的识别速度。

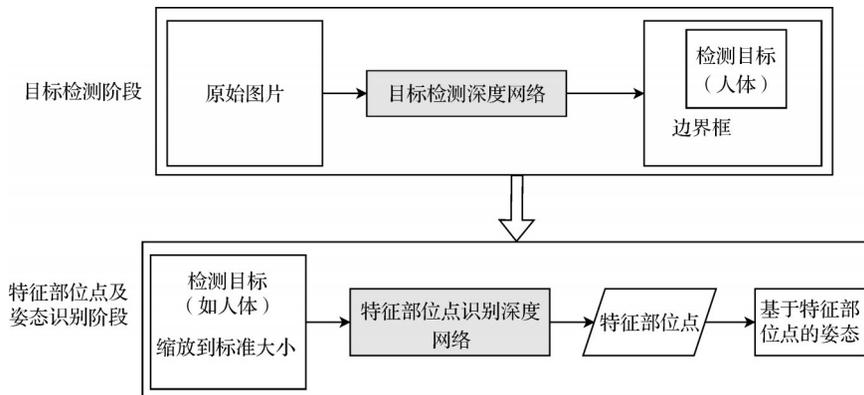


图2 自上而下的姿态识别方法

Fig.2 Top-to-down position recognition methods

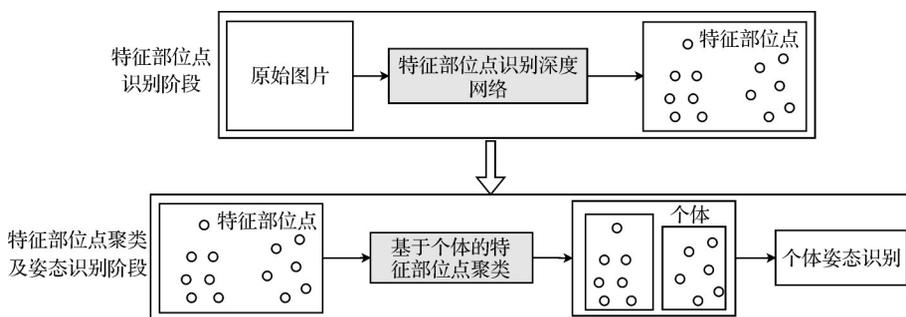


图3 自下而上的姿态识别方法

Fig.3 Bottom-to-up position recognition methods

Cao等^[34-35]提出的OpenPose模型在姿态识别领域获得了广泛的应用。基于非参数表达的部位关联场(part affinity fields, PAFs)算法,采用了自下而上方式来识别图片中每个人的姿态。首先识别出图片中所有人的关键部位点,然后根据关键部位点间的关联程度确定哪些关键点是属于同一个体,无论图片中存在多少人,都可以在具有较高准确率的情况下获得实时检测效果。Chen等^[36]采用了自上而下

的方法,首先检测出图像中的每个个体,并用边界框标记出来,然后再针对个体检测其关键部位点。Wei等^[37]提出了卷积姿态机(convolutional pose machines, CPM),其同样采用了自上而下的关键点检测算法,需要结合人体检测算法来进行人体姿态识别。Fang等^[38]基于自上而下的方式提出了针对多人姿态的检测方法,能够在人体检测边界框存在误差的情况下准确地检测出人体姿态。Newell等^[39]采用了

重复利用从下到上及从上到下的检测方式,利用堆叠沙漏网络进行人体姿态估计。Sun等^[40]利用输入图像的高分辨率表征,并与不同层级的低分辨率子网络相结合,明显提升了人体关键部位点检测的准确率。

行为识别多是通过检测人体在一定时间段内姿态或动作变化的情况来评估人的行为。行为识别与姿态识别十分类似,但却是2种不同的任务^[41],通常需要采用视频的方式引入时间维度的信息以进行行为识别。Luvizon等^[41]提出了一种姿态与行为同时识别的多任务框架,基于静态图片进行2D/3D姿态估计,同时基于视频序列进行人体行为识别,实现了通过单一的网络结构解决2种类型问题的目标。Chu等^[42]在提取个体特征的基础上引入时空注意力机制(spatial-temporal attention mechanism, STAM),以解决目标追踪过程中由于遮挡及个体间相互作用产生的漂移问题。在个体追踪及行为分析方面,由于个体的动作行为是一个时序过程,通常需要考虑个体的时序特征。因此,针对时间序列的深度循环神经网络(recurrent neural network, RNN)被广泛应用于行为特征的自动提取中。在人的运动轨迹追踪过程中, Sadeghian等^[43]提出了基于RNN的多目标追踪模型,利用人在移动中的时序变化特征、人与人之间的相互关系及所提取的个体表现特征来实现运动追踪,同时解决运动中的遮挡问题。Jaouedi等^[44]为了更好地对人的行为进行分类,同样采取了运动追踪的方式,采用混合高斯模型及卡尔曼滤波来检测移动的人体,通过门控递归神经网络(gated recurrent neural networks, GRNN)来提取每帧的特征并预测人的行为。然而,这些技术在智能养猪领域的应用探索中,仍然存在一些挑战。由于定位目标是姿态识别和行为分析的前提,在实验室场景中很多模型都能有效通过图像的目标检测或目标分割来实现,然而在实际的现代化养殖场景中,密集的猪只个体、有限的拍摄位置视角等因素都难以满足如此严苛的视觉环境条件。

2 猪只个体身份与行为识别方法的研究进展

2.1 猪只个体身份识别方法

生猪养殖大多以群体方式,准确地识别猪只个体身份是检测猪只行为的前提。传统的猪只识别,主要采用标记、耳缺或射频识别(radio frequency

identification, RFID)耳标的方式。采用耳缺的方式,是通过缺口的编码来标识猪只身份,这种方式一方面不利于动物福利,另一方面需要人工读取。采用RFID方式可以利用自动读取设备自动获得身份的ID,但目前没有获得广泛应用。原因是:RFID标签成本较高,安装耗时;同时RFID标签的复用造成了猪只ID全局不唯一的情况;更为关键的是,对于群养猪只,当多个猪只靠近RFID读卡器设备时,难以真正做到猪只个体身份的识别;此外,RFID标签的可分离特征,使猪肉溯源时容易被伪造。因此,采用基于生物特征的图像及视频方式来获取猪只身份,实时高效,不可伪造,且不需要在猪只身上使用额外设备就能够识别猪只身份并检测其行为,是未来猪只身份识别与行为检测的发展趋势。

通过视频来识别猪只身份最简单的方式是在猪只身上做标记,然后通过识别视频中的标记来获取猪只身份^[45],但这种方式存在标记易变、易掉落、易被污染的问题。同时标记的编码位数有限,只适合在小范围的猪群中用于身份的标识。因此,与人脸识别一样,采用深度卷积神经网络提取猪只的生物表现特征,是猪只个体身份识别的一种更有效方法。

利用生物特征对猪只个体身份进行识别,通常分为2个阶段:首先通过猪脸或身体检测模型进行检测,然后用矩形边界框标注出猪脸或目标部位的位置,再通过深度网络提取猪脸或身体的特征向量,最后识别出猪只个体的身份。Hansen等^[46]选取猪场中10头不同猪只,参考人脸识别构建卷积神经网络进行猪只个体身份识别。结果表明,算法所提取的特征大多源自猪只面部,如鼻子、眼睛等部位,1553张图像统计的个体猪只识别准确率为96.7%。Wang等^[47]采用了多个深度神经网络对猪只特征进行提取,并整合这些特征进行猪只身份识别。这种方式可有效提高识别准确度,但是由于利用了多个深度神经网络,会加大算法对计算资源的需求。Marsot等^[48]提出了一种具有相对较低成本且易扩展的图像与深度卷积网络的猪只识别方法:首先,采用在人脸检测中常用的2个Haar特征级联分类器^[49]检测出猪只的面部与眼睛,然后利用1个浅层卷积神经网络提取高质量图片,另1个深层卷积神经网络进行猪脸识别,最终在320张图片中识别准确率达83%。Wang等^[50]提出了一个基于深度神经网络的轻型猪脸识别模型,该模型包含了2个阶段,第一阶段是基于EfficientDet-D0模型^[51]的猪脸检测,采用数据集采样技

术将猪脸检测的平均精度从90.7%提高到99.1%;第二阶段是猪脸分类,采用ResNet-18等8个分类模型,并基于triplet margin loss函数做了方法改进,通过集成两阶段模型对28头猪的脸部识别的平均精度值为91.35%,与采用单阶段法训练的EfficientDet-D0网络相比,平均精度值提高了28%。Wang等^[52]采用了级联猪脸检测与猪脸识别的深度学习模型,首先从图像中检测出猪脸,然后基于检测出的猪脸进行猪只个体身份识别。采用包含46头猪的数据集,其中41头猪用来构建训练集,5头猪用来构建测试集来验证模型的准确性,级联模型在3个数据集上的准确率分别达到99.38%、98.96%和97.66%。基于改进的YOLOv3及YOLOv4进行猪只身份识别^[53-54],其中文献[53]是针对猪只整体而不是猪脸进行的身份识别,当IoU阈值为0.5,分类概率阈值为0.1时,YOLOv3_DB_SPP模型的平均精度比YOLOv3模型提高了9.87%。改进的YOLOv4模型^[54]提高了个体猪只的识别精度和速度,当阈值为0.5时,测试集的平均准确率达到98.12%,召回率达到95%, F_1 分数为96%,平均FPS为34.3 ms,平均交并比为83.91%。YOLOv3_DB_SPP模型^[55]针对猪只识别领域中如遮挡、小目标样本识别等方面也做了很好的深入研究。张建龙等^[56]研究了生长过程中肉猪脸部的变化对模型识别准确率的影响,重点分析了DenseNet201、MobileNetv3_small、SeNet154、Xception这4种模型,结果表明SeNet154模型的效果最佳,对验证集的识别准确率可达98.80%,随着猪只生长带来的脸部变化及环境条件变化,模型的识别准确率逐天下降,建议猪脸检测时至少使用前4d的数据以更新猪脸识别模型。综上,在通用型计算机视觉与深度学习模型的基础上,这些模型针对猪只的特征进行相应的改进或采用迁移学习的算法,进一步提升了猪只个体身份与行为识别的准确率。然而,在从人到动物的领域转换过程中,目前很少有针对猪只身份及行为特征的基于计算机视觉的原创性深度学习模型的研究。

2.2 猪只特定行为识别方法

对于猪只行为的识别,已有研究工作主要分为两大类,一类是针对某种具体行为的特定行为识别方法,另一类则是对于猪只多类型行为的通用行为识别方法。

由于采食及饮水行为是判断猪只健康状态的重要指标,有大量学者展开了相关的研究。Yang等^[4]

基于Faster R-CNN网络及饲喂区域的像素占有率提出了一种采食行为识别算法,首先识别出猪只头部,然后根据头部在采食区域的移动状况来判断采食行为。对猪只身份识别采用背部字母标记的方式,背部字母检测法效率虽高但存在易退色、易受污染的问题。采食行为对于仔猪的健康状况尤其重要, Kim等^[57]针对断奶仔猪采食行为进行识别,采用了YOLOv3与YOLOv4模型,主要关注于检测猪只在料槽旁的采食行为。与传统方法(先检测目标猪只,再检测不同的猪只行为)不同的是,所提出的方法是将这2个步骤合并以提升检测速度。同样针对仔猪, Chen等^[58]提出了一种基于视频的深度学习算法来识别采食行为,并能够确定每头猪只个体的采食时间,采用Xception卷积神经网络架构^[59]来提取空间特征,然后将这些空间特征输入到长短期记忆(long short-term memory, LSTM)框架进一步提取时空特征,通过二分类的方式来判断猪只是否正在进食,并提出了一种基于最大熵分割、HSV(Hue, saturation and value)颜色的图像处理算法,采用空间变换和模板匹配来计算头部的圆度、头部与采食子部位的比例、头部运动的累积像素以及头部到猪背上的编号距离,以确定每头猪的身份和喂食时间。同样,李菊霞等^[60]针对猪只饮食行为的自动化检测展开研究,利用标注的猪只行为图片训练YOLOv4网络进行特定行为的识别,结果表明,在不同的视角、遮挡、时间、光照等条件下均能够准确识别,但对检测模型本身并没有做相关的修改。因此,采用成熟的深度学习模型,通过标注猪只行为的图片进行训练,能够较准确地检测出猪只的行为。Zhu等^[61]首先识别猪栏中7头猪的身份,然后通过测量图像中猪只与水龙头的距离来判定饮水活动。Alameer等^[62]同样是基于深度学习技术对猪只的饮水行为进行识别,并通过饮水行为判断猪只的健康状况。

此外,对于攻击、趴跨、哺乳等特定行为,学者都开展了相应的智能化方法的研究。如Chen等^[63]利用Kinect及能量模型识别猪群中猪只的攻击行为,并提取其特征。Yang等^[64]利用图像分割技术,通过母猪和仔猪的姿态与空间位置关系及基于光流帧的时序运动特征识别母猪趴跨行为。李丹等^[65-66]则是针对趴跨这一特定行为进行识别,利用猪只俯视图,基于Mask R-CNN网络进行图像分割,并采取二分类的方式来判断猪只是否出现了爬跨行为。

2.3 猪只多类型行为识别方法

猪只多类型行为的识别方法具有更好的通用性,可以识别出猪只常见的各类动作与行为,是当前猪只行为识别的一个重要研究方向。

董力中等^[67]提出的猪只行为识别方法充分利用了猪只的姿态与时序特征,利用YOLOv5进行猪只检测,利用OpenPose算法进行猪只姿态的估计,对猪只身体的20个关键部位如鼻、耳朵、肩、膝等及其相关联的26个关键点数据进行分析,并对猪只骨架进行提取,同时引入时序信息,利用时空图卷积网络对猪只的站、走、卧3种行为进行识别。涂淑琴等^[68]同样基于YOLOv5深度模型并结合改进的DeepSort行为追踪算法,对猪只行为进行识别与追踪,改善了猪只重叠与遮挡造成的猪只身份编号频繁跳变的问题。首先,基于YOLOv5模型采用分类的方式检测出猪只的躺卧、站立、饮食及其他行为等4种不同类别的猪只行为,然后基于猪只当前行为类别,利用改进的DeepSort追踪算法实现在猪只跟踪过程中行为的识别。高云等^[69]基于2D卷积网络的识别方法,针对单张图片对猪只行为识别不准确的问题,通过在当前图像的基础上添加一定数目的后续视频图像,实现了时间维度信息的卷积,以此构建3D卷积神经网络,进而对撕咬等4种大类、咬耳等7种小类侵略性行为及吃食、饮水、休息等非侵略性行为进行识别。李光昌等^[70]将OpenPose人体姿态估计算法进行改造,以检测猪只姿态,首先,采用OpenPose算法得到猪只身体的关键部位点,然后通过关键部位点之间的位置关系来获得猪只个体的行为特征,再采用KNN(k -nearest neighbor)算法进行分类,最后对猪只站、趴、卧、坐4种典型姿势进行识别。杨秋妹^[71]先对猪只进行检测,再采用分类的方式,基于猪只当前行为类别预测识别猪只行为。首先利用GoogleNet深度网络对猪只个体进行姿态分类的方式识别出躺、趴、站3种姿态,然后采用基于ECO-HC的跟踪算法实现猪只运动估计的持续跟踪,再在识别出姿态的基础上结合猪只中心的位置变化以识别猪只的休息、走动、静站等行为,最后结合时间信息、当前猪只所在猪舍的功能、猪只间的位置关系变化等附加信息实现对猪只饮水、进食、爬跨等行为的识别,并能够通过分析视频各帧之间的关联,对猪群整体的活跃度进行评估。对于猪只的个体识别,同样是采用在猪只背部打上字母标记,然后识别字母标记,进而识别猪只个体的方式。

3 小 结

综上,精准定位到猪只个体的行为才符合未来规模化猪场的养殖需求。在猪场监管中,基于计算机视觉及深度学习的相关模型与算法在目标检测与识别上已取得很好的成果,归纳如下:

(1)猪只个体身份与行为识别,充分借鉴了人的身份识别领域相关的计算机视觉与深度学习的模型与方法。同时,各类新型的卷积神经网络能够快速地在猪只个体身份与行为识别领域得到应用。

(2)在计算机视觉与深度学习模型的基础上,针对猪只的特征进行相应的改进,进一步提升了猪只个体身份与行为识别的准确率。

(3)越来越多的研究表明猪只个体身份识别与行为识别之间存在紧密关系,基于猪只个体身份的行为识别能够准确定位到特定的猪只,为精准化养殖奠定了基础。

(4)通用型的猪只行为识别模型能够识别各种类型的猪只行为,适用于各类场景,已成为当前研究的一个热点问题。

4 展 望

基于计算机视觉与深度学习网络的猪只个体身份与行为识别研究,目前已取得了一定进展,因其独有特征,仍存在一些关键问题需要解决,具体如下:

(1)猪只运动随意,特征部位如面部等图像采集困难,且容易受到环境污染与挤压遮挡。因此,如何在猪只运动不可控及关键特征部位受到污染的情况下,准确提取其身份及行为特征,是影响动物身份准确识别与行为精准检测的关键问题。

(2)当前的猪只个体身份与行为识别方法所采用的人工神经网络模型,大多来自于人的身份及行为识别领域。很少有针对猪只身份及行为特征的基于计算机视觉的原创性深度学习模型的研究。从人转变到动物的领域,通过对于模型的改进或采用迁移学习的算法,能够在一定程度上识别出猪只身份与行为。但是,这些模型并没有对动物本身的习性与特征展开有针对性的设计,因此,针对猪只特征而设计有针对性的原创性深度学习模型意义重大。

(3)对于生猪的智能化养殖,个体身份识别与行为识别往往需要同时进行,而目前的方法大多是分别采用2个或多个深度神经网络,这无疑会带来额外的计算冗余。无论是身份识别还是行为识别,都是从图像或视频中提取生物属性特征,然后进行身份

或行为的识别。因此,构建多任务检测网络,融合身份、姿态及行为的检测与识别功能,对于提升算法的性能,降低实际应用中部署的成本具有重要的价值。

(4)通用型的猪只识别模型是未来重要的研究方向之一。首先识别猪只的各类基本姿态及动作,然后根据对这些基本的姿态及动作进行时序关联分析等,进而识别出猪只的行为。通用型的识别模型的建立,可以满足不同场景下猪只行为的识别问题,具有一定的通用性,是未来猪只行为识别的关键。

(5)采用边缘计算进行猪只个体身份与行为的识别是深度模型部署与实现的关键。如何将开发出的猪只个体身份与行为识别模型部署在实际的场景中并能够真正解决应用问题,是将研究成果进行产业转化的关键。目前随着物联网及边缘计算的快速发展,将算法模型部署在边缘端,通过边缘计算的能力来实现深度学习模型的推理是一个主流方向。因此,基于边缘计算的猪只身份与行为识别模型的部署与应用是未来研究的关键,该问题的解决对于提升识别的效率具有重大意义。

解决上述问题,对于猪只身份、姿态与行为识别的准确度的提升,生猪精准养殖的自动化水平的提高,以及畜牧业智能化的发展具有重要的理论与实践意义。

参考文献 References

- [1] YANG Q M, XIAO D Q. A review of video-based pig behavior recognition[J/OL]. *Applied animal behaviour science*, 2020, 233: 105146 [2022-10-31]. <https://doi.org/10.1016/j.applanim.2020.105146>.
- [2] 谭辉磊. 基于机器视觉的猪个体身份和饮水行为识别方法[D]. 镇江: 江苏大学, 2017. TAN H L. Recognition method of identification and drinking behavior for individual pigs based on machine vision[D]. Zhenjiang: Jiangsu University, 2017 (in Chinese with English abstract).
- [3] 史怀, 陈峥, 潘志针. 基于机器视觉的猪个体身份和行为识别[J]. *福建畜牧兽医*, 2021, 43(1): 27-30. SHI H, CHEN Z, PAN Z Z. Identification of pig individual identity and behavior based on machine vision[J]. *Fujian journal of animal husbandry and veterinary medicine*, 2021, 43(1): 27-30 (in Chinese with English abstract).
- [4] YANG Q M, XIAO D Q, LIN S C. Feeding behavior recognition for group-housed pigs with the Faster R-CNN[J]. *Computers and electronics in agriculture*, 2018, 155: 453-460.
- [5] 马娜, 徐苗. 基于卷积神经网络的猪只个体身份识别研究[J]. *计算机时代*, 2022(4): 51-54. MA N, XU M. Research on pig individual identity based on convolutional neural networks[J]. *Computer era*, 2022(4): 51-54 (in Chinese with English abstract).
- [6] OUYANG W L, ZENG X Y, WANG X G, et al. DeepID-net: object detection with deformable part based convolutional neural networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(7): 1320-1334.
- [7] DIBA A L, SHARMA V, PAZANDEH A, et al. Weakly supervised cascaded convolutional networks[C]//2017 IEEE conference on computer vision and pattern recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 5131-5139.
- [8] DOULAMIS N, VOULODIMOS A. FAST-MDL: fast adaptive supervised training of multi-layered deep learning models for consistent object tracking and classification[C]//2016 IEEE international conference on imaging systems and techniques (IST), October 4-6, 2016, Chania, Greece. New York: IEEE, 2016: 318-323.
- [9] LIN L, WANG K Z, ZUO W M, et al. A deep structured model with radius-margin bound for 3D human activity recognition[J]. *International journal of computer vision*, 2016, 118(2): 256-273.
- [10] TOSHEV A, SZEGEDY C. DeepPose: human pose estimation via deep neural networks[C]//2014 IEEE conference on computer vision and pattern recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1653-1660.
- [11] NOH H, HONG S, HAN B. Learning deconvolution network for semantic segmentation[C]//2015 IEEE international conference on computer vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2016: 1520-1528.
- [12] VOULODIMOS A, DOULAMIS N, DOULAMIS A, et al. Deep learning for computer vision: a brief review[J/OL]. *Computational intelligence and neuroscience*, 2018: 7068349 [2022-10-31]. <https://doi.org/10.1155/2018/7068349>.
- [13] BASTIEN F, LAMBLIN P, PASCANU R, et al. Theano: new features and speed improvements[J]. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 2012, 1211: 5590-5598.
- [14] KAKANI V, NGUYEN V H, KUMAR B P, et al. A critical review on computer vision and artificial intelligence in food industry[J/OL]. *Journal of agriculture and food research*, 2020, 2: 100033 [2022-10-31]. <https://doi.org/10.1016/j.jafr.2020.100033>.
- [15] 安小松, 宋竹平, 梁千月, 等. 基于CNN-Transformer的视觉缺陷柑橘分选方法[J]. *华中农业大学学报*, 2022, 41(4): 158-169. AN X S, SONG Z H, LIANG Q Y, et al. A CNN-Transformer-based method for sorting citrus with visual defects[J]. *Journal of Huazhong Agricultural University*, 2022, 41(4): 158-169 (in Chinese with English abstract).
- [16] HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision[M]. 2nd ed. Cambridge, UK: Cambridge University Press, 2003.
- [17] VEDALDI A, FULKERSON B. Vifeat: an open and portable library of computer vision algorithms[C]//Proceedings of the 18th ACM international conference on Multimedia. New York: ACM, 2010: 1469-1472.
- [18] 顾伟, 王巧华, 李庆旭, 等. 基于改进SSD的棉种破损检测[J]. *华中农业大学学报*, 2021, 40(3): 278-285. GU W, WANG Q H, LI Q X, et al. Improved SSD based detection of damaged cot-

- tonseed [J]. *Journal of Huazhong Agricultural University*, 2022, 40(3): 278-285 (in Chinese with English abstract).
- [19] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [DB/OL]. arXiv, 2020: 1704.04861 [2022-10-31]. <https://doi.org/10.48550/arXiv.1704.04861>.
- [20] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [21] REDMON J, FARHADI A. YOLOv3: an incremental improvement [DB/OL]. arXiv, 2018: 1804.02767 [2022-10-31]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [22] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [DB/OL]. arXiv, 2020: 2004.10934 [2022-10-31]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [23] ZHANG K P, ZHANG Z P, LI Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. *IEEE signal processing letters*, 2016, 23(10): 1499-1503.
- [24] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE conference on computer vision and pattern recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 580-587.
- [25] GIRSHICK R. Fast R-CNN [C]//2015 IEEE international conference on computer vision (ICCV). December 7-13, 2015, Santiago, Chile. New York: IEEE, 2016: 1440-1448.
- [26] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-1149.
- [27] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//2017 IEEE international conference on computer vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2980-2988.
- [28] WANG X L, KONG T, SHEN C H, et al. SOLO: segmenting objects by locations [M]//Computer vision -ECCV 2020. Cham: Springer International Publishing, 2020: 649-665.
- [29] TAIGMAN Y, YANG M, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification [C]//2014 IEEE conference on computer vision and pattern recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1701-1708.
- [30] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering [C]//2015 IEEE conference on computer vision and pattern recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 815-823.
- [31] ZHOU E J, CAO Z M, YIN Q. Naive-deep face recognition: touching the limit of LFW benchmark or not? [DB/OL]. arXiv, 2015: 1501.04690 [2022-10-31]. <https://doi.org/10.48550/arXiv.1501.04690>.
- [32] LIU J T, DENG Y F, BAI T, et al. Targeting ultimate accuracy: face recognition via deep embedding [DB/OL]. arXiv, 2015: 1506.07310 [2022-10-31]. <https://doi.org/10.48550/arXiv.1506.07310>.
- [33] LIU W Y, WEN Y D, YU Z D, et al. SphereFace: deep hypersphere embedding for face recognition [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6738-6746.
- [34] CAO Z, SIMON T, WEI S H, et al. Realtime multi-person 2D pose estimation using part affinity fields [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1302-1310.
- [35] CAO Z, HIDALGO G, SIMON T, et al. OpenPose: realtime multi-person 2D pose estimation using part affinity fields [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2021, 43(1): 172-186.
- [36] CHEN Y L, WANG Z C, PENG Y X, et al. Cascaded pyramid network for multi-person pose estimation [C]//2018 IEEE/CVF conference on computer vision and pattern recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7103-7112.
- [37] WEI S H, RAMAKRISHNA V, KANADE T, et al. Convolutional pose machines [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 4724-4732.
- [38] FANG H S, XIE S Q, TAI Y W, et al. RMPE: regional multi-person pose estimation [C]//2017 IEEE international conference on computer vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2353-2362.
- [39] NEWELL A, YANG K Y, DENG J. Stacked hourglass networks for human pose estimation [M]//Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016: 483-499.
- [40] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation [C]//2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2020: 5686-5696.
- [41] LUVIZON D C, PICARD D, TABIA H. 2D/3D pose estimation and action recognition using multitask deep learning [C]//2018 IEEE/CVF conference on computer vision and pattern recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 5137-5146.
- [42] CHU Q, OUYANG W L, LI H S, et al. Online multi-object tracking using CNN-based single object tracker with spatial-temporal attention mechanism [C]//2017 IEEE international conference on computer vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 4846-4855.
- [43] SADEGHIAN A, ALAHI A, SAVARESE S. Tracking the untrackable: learning to track multiple cues with long-term depen-

- dencies[C]//2017 IEEE international conference on computer vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 300-311.
- [44] JAOUEDI N, BOUJNAH N, BOUHLEL M S. A new hybrid deep learning model for human action recognition[J]. Journal of King Saud University-computer and information sciences, 2020, 32(4): 447-453.
- [45] KASHIHA M, BAHR C, OTT S, et al. Automatic identification of marked pigs in a pen using image pattern recognition[J]. Computers and electronics in agriculture, 2013, 93: 111-120.
- [46] HANSEN M F, SMITH M L, SMITH L N, et al. Towards on-farm pig face recognition using convolutional neural networks[J]. Computers in industry, 2018, 98: 145-152.
- [47] WANG J Z, LIU A Z, XIAO J. Video-based pig recognition with feature-integrated transfer learning [M]//PATO J N, MILLETT L I. Biometric recognition. Cham: Springer International Publishing, 2018: 620-631.
- [48] MARSOT M, MEI J Q, SHAN X C, et al. An adaptive pig face recognition approach using convolutional neural networks [J/OL]. Computers and electronics in agriculture, 2020, 173: 105386 [2022-10-31]. <https://doi.org/10.1016/j.compag.2020.105386>.
- [49] SHARIFARA A, RAHIM M, ANISI Y. A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection[J]. Proceedings-ISAFAST, 2014, 2015: 73-78.
- [50] WANG Z Y, LIU T H. Two-stage method based on triplet margin loss for pig face recognition[J/OL]. Computers and electronics in agriculture, 2022, 194: 106737 [2022-10-31]. <https://doi.org/10.1016/j.compag.2022.106737>.
- [51] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 10778-10787.
- [52] WANG R, SHI Z F, LI Q F, et al. Pig face recognition model based on a cascaded network[J]. Applied engineering in agriculture, 2021, 37(5): 879-890.
- [53] LI G B, JIAO J, SHI G L, et al. Fast recognition of pig faces based on improved YOLOv3[J/OL]. Journal of physics: conference series, 2022, 2171(1): 012005 [2022-10-31]. <https://doi.org/10.1088/1742-6596/2171/1/012005>.
- [54] LI S D, KANG X, FENG Y K, et al. Detection method for individual pig based on improved YOLOv4 convolutional neural network[C]//DSIT 2021: 2021 4th international conference on data science and information technology, July 23-25, 2021, Shanghai, China. New York: ACM, 2021: 231-235.
- [55] 何屿彤, 李斌, 张锋, 等. 基于改进 YOLOv3 的猪脸识别[J]. 中国农业大学学报, 2021, 26(3): 53-62. HE Y T, LI B, ZHANG F, et al. Pig face recognition based on improved YOLOv3[J]. Journal of China Agricultural University, 2021, 26(3): 53-62 (in Chinese with English abstract).
- [56] 张建龙, 周康, 庄晏榕, 等. 育肥猪生长过程中脸部变化对识别模型准确率的影响[J]. 中国农业大学学报, 2021, 26(6): 180-186. ZHANG J L, ZHOU K, ZHUANG Y R, et al. Effect of facial changes on the accuracy of the recognition model during the growth of finishing pigs[J]. Journal of China Agricultural University, 2021, 26(6): 180-186 (in Chinese with English abstract).
- [57] KIM M, CHOI Y, LEE J N, et al. A deep learning-based approach for feeding behavior recognition of weanling pigs[J]. Journal of animal science and technology, 2021, 63(6): 1453-1463.
- [58] CHEN C, ZHU W X, STEIBEL J, et al. Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method[J]. Computers and electronics in agriculture, 2020, 176: 105642 [2022-10-31]. <https://doi.org/10.1016/j.compag.2020.105642>.
- [59] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1800-1807.
- [60] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251-256. LI J X, LI Y W, NIU F, et al. Pig diet behavior detection method based on YOLOv4[J]. Transactions of the CSAM, 2021, 52(3): 251-256 (in Chinese with English abstract).
- [61] ZHU W X, GUO Y Z, JIAO P P, et al. Recognition and drinking behaviour analysis of individual pigs based on machine vision[J]. Livestock science, 2017, 205: 129-136.
- [62] ALAMEER A, KYRIAZAKIS I, BACARDIT J. Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs [J/OL]. Scientific reports, 2020, 10: 13665 [2022-10-31]. <https://doi.org/10.1038/s41598-020-70688-6>.
- [63] CHEN C, ZHU W, GUO Y, et al. A kinetic energy model based on machine vision for recognition of aggressive behaviours among group-housed pigs [J]. Livestock science, 2018, 218: 70-78.
- [64] YANG A Q, HUANG H S, ZHU X M, et al. Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features[J]. Biosystems engineering, 2018, 175: 133-145.
- [65] 李丹, 张凯锋, 李行健, 等. 基于 Mask R-CNN 的猪只爬跨行为识别[J]. 农业机械学报, 2019, 50(S1): 261-266, 275. LI D, ZHANG K F, LI X J, et al. Mounting behavior recognition for pigs based on mask R-CNN [J]. Transactions of the CSAM, 2019, 50(S1): 261-266, 275 (in Chinese with English abstract).
- [66] LI D, CHEN Y F, ZHANG K F, et al. Mounting behaviour recognition for pigs based on deep learning[J/OL]. Sensors (Basel, Switzerland), 2019, 19(22): 4924 [2022-10-31]. <https://doi.org/10.3390/s19224924>.
- [67] 董力中, 孟祥宝, 潘明, 等. 基于姿态与时序特征的猪只行为识别方法[J]. 农业工程学报, 2022, 38(5): 148-157. DONG L Z, MENG X B, PAN M, et al. Recognizing pig behavior on posture and temporal features using computer vision[J]. Transactions of the CSAE, 2022, 38(5): 148-157 (in Chinese with English abstract).
- [68] 涂淑琴, 刘晓龙, 梁云, 等. 基于改进 DeepSORT 的群养生猪行为识别与跟踪方法[J]. 农业机械学报, 2022, 53(8): 345-352.

- TU S Q, LIU X L, LIANG Y, et al. Behavior recognition and tracking method of group-housed pigs based on improved Deep-SORT algorithm[J]. Transactions of the CSAM, 2022, 53(8): 345-352 (in Chinese with English abstract).
- [69] 高云, 陈斌, 廖慧敏, 等. 群养猪侵略性行为的深度学习识别方法[J]. 农业工程学报, 2019, 35(23): 192-200. GAO Y, CHEN B, LIAO H M, et al. Recognition method for aggressive behavior of group pigs based on deep learning[J]. Transactions of the CSAE, 2019, 35(23): 192-200 (in Chinese with English abstract).
- [70] 李光昌, 刘飞飞, 李嘉豪. 基于改进OpenPose算法的猪只行为识别方法[J]. 河南农业大学学报, 2022, 56(3): 460-470. LI G C, LIU F F, LI J H. Pig behavior recognition method based on improved OpenPose algorithm[J]. Journal of Henan Agricultural University, 2022, 56(3): 460-470 (in Chinese with English abstract).
- [71] 杨秋妹. 基于现场视频的猪只行为识别[D]. 广州: 华南农业大学, 2019. YANG Q M. Pig behavior recognition based on live video [D]. Guangzhou: South China Agricultural University, 2019 (in Chinese with English abstract).

Progress of computer vision and deep learning methods for pig's identity and behavior recognition

LIU Feng^{1,2,3}, WU Wenjie¹, LIU Xiaolei^{2,3,4}, WANG Xinran¹,
FANG Yaping¹, LI Guoliang^{1,2,3}, DU Xiaoyong^{1,2,3,4}

1. College of Informatics, Huazhong Agricultural University/Key Laboratory of Intelligent Technology in Animal Husbandry, Ministry of Agriculture and Rural Affairs/
Engineering Research Center of Smart Agricultural Technology, Ministry of Education/
Hubei Province Research Center of Engineering Technology of Agricultural
Big Data, Wuhan 430070, China;

2. Shenzhen Institute of Nutrition and Health, Huazhong Agricultural University, Shenzhen 518000, China;

3. Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences/
Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Shenzhen 518000, China;

4. Key Laboratory of Swine Genetics and Breeding, Ministry of Agriculture, Wuhan 430070, China

Abstract It is an important study direction in the area of smart farming to explore the combination of new progress in the field of pig farming with artificial intelligence. Among them, how to automatically identify the individual identity and behavior of pigs is a key and hot issue to be solved in the current pig breeding industry. This article summarizes the existed methods of using deep neural networks to identify the individual identity and behavior of pigs based on the progress of computer vision and deep learning models in human recognition. The problems in the existed methods are analyzed, and the key study directions in the future are proposed. Five aspects urgently needed to be developed are as follows: (1) the methods of accurately extracting the features of pig's identity and behavior under the conditions that pig's behavior cannot be controlled and the key parts of pig's body are often contaminated; (2) the deep learning models based on computer vision that dedicate for pigs to recognize the identity and behavior based on the specific features of pigs; (3) the studies on multi-task deep learning models that can recognize pig's identity and behavior simultaneously; (4) the studies on general-purpose pig behavior recognition methods based on basic postures and movements that are applicable to multiple scenarios; (5) the studies on the deployment methods of pig identification and behavior recognition based on edge computing.

Keywords identity recognition; behavior recognition; deep learning; computer vision; smart farming; pig face recognition; automatic monitoring

(责任编辑: 赵琳琳)