

基于 YOLO v3 与图结构模型的群养猪只头尾辨别方法

李 泊 沈明霞 刘龙申 陆明洲 孙玉文  
(南京农业大学工学院, 南京 210031)

**摘要:** 在利用视频监控技术对群养猪只进行自动行为监测时,对猪只准确定位并辨别其头尾位置对提高监测水平至关重要,基于此提出一种基于 YOLO v3 (You only look once v3)模型与图结构模型(Pictorial structure models)的猪只头尾辨别方法。首先,利用基于深度卷积神经网络的 YOLO v3 目标检测模型,训练猪只整体及其头部和尾部 3 类目标的检测器,从而在输入图像中获得猪只整体及头尾部所有的检测结果;然后,引入图结构模型,描述猪只的头尾结构特征,对每个猪只整体检测矩形框内的头尾部位组合计算匹配得分,选择最优的部位组合方式;对部分部位漏检的情况,采取阈值分割与前景椭圆拟合的方法,根据椭圆长轴推理出缺失部位。在实际猪场环境下,通过俯拍获得猪舍监控视频,建立了图像数据集,并进行了检测实验。实验结果表明,与直接利用 YOLO v3 模型相比,本文方法对头尾定位的精确率和召回率均有一定提高。本文方法对猪只头尾辨别精确率达到 96.22%,与其他方法相比具有明显优势。

**关键词:** 猪只; 行为识别; 头尾辨别; 深度卷积神经网络; 图结构模型

中图分类号: TP391.41      文献标识码: A      文章编号: 1000-1298(2020)07-0044-08      OSID: 

Head and Tail Identification Method for Group-housed Pigs  
Based on YOLO v3 and Pictorial Structure Model

LI Bo SHEN Mingxia LIU Longshen LU Mingzhou SUN Yuwen  
(College of Engineering, Nanjing Agricultural University, Nanjing 210031, China)

**Abstract:** For automatic behavior monitoring of group-housed pigs in video surveillance, pig head/tail identification has important significance to improve the level of behavior recognition. A head-tail recognition algorithm was proposed based on YOLO v3 (You only look once v3) and pictorial structure models. Firstly, the object detectors of three categories, i. e. , pigs, heads and tails, were trained with YOLO v3, which was a general object detection model based on deep convolutional neural networks. In this way, bounding boxes of pigs, heads and tails can be detected from the input image. Next, pictorial structure models were introduced to describe structural characteristics of heads and tails for pigs. For each detected bounding box of pigs, scores of all possible head-tail combinations were computed with the established pictorial structure model to choose the optimal part configuration. When a head or tail was missed in the pig bounding box, a part inference method based on threshold segmentation was utilized to estimate the missing part according to the major axis of the fitted ellipse. In experiments, an image dataset was constructed from a top-view surveillance video of group-housed pigs. Experimental results demonstrated that via the proposed method, the precision and recall of part localization were improved compared with results of YOLO v3. Moreover, the head/tail identification accuracy reached 96.22%, which obviously outperformed other methods based on intersection of bounding boxes and generalized Hough clustering. As a result, the proposed method can effectively detect pigs and distinguish their heads/tails in images of group-housed pigs without excessive limitations on environments.

**Key words:** pig; behavior recognition; head/tail identification; deep convolutional neural networks; pictorial structure model

## 0 引言

智能视频监控技术在生猪养殖中的应用日趋广泛,利用计算机视觉技术提高生猪行为监测水平已成为农业工程领域的研究热点之一<sup>[1-3]</sup>。要实现视频图像中群养猪只的自动行为监测,准确的猪只定位是系统必不可少的环节<sup>[4-6]</sup>。随着研究的深入,仅获得整体猪只的位置已不能完全满足精准监测的需求。例如在猪只饮水、采食、爬跨等行为的监测识别过程中,相比于整体猪只的位置,利用头部或尾部等部位信息明显更加合理<sup>[7-11]</sup>。因此,准确获取猪只的头部和尾部等关键部位信息对于提高猪只的行为监测水平具有重要意义。

近年来,对于图像中的猪只检测问题国内外有许多深入研究,旨在不断提高猪只检测的精度<sup>[12-15]</sup>。而关于猪只头尾辨别问题的研究相对较少,并且大部分情况下仅将此问题作为整个系统的一个简单操作进行处理,因此采用方法都较为基础。KASHIHA 等<sup>[8]</sup>通过分析猪只轮廓各点到质心距离曲线的整体趋势来判断头部位置。NASIRAHMADI 等<sup>[9]</sup>利用拟合椭圆的长轴与猪体轮廓的交点判断头尾位置,但并未分辨出头部和尾部。杨心等<sup>[16]</sup>在通过椭圆拟合及猪体身长截取出猪体两端的轮廓后,利用广义 Hough 聚类的方法识别轮廓属于头部或尾部。高云等<sup>[17]</sup>在此方法的基础上增加了对轮廓圆度的考虑,进一步提高了头尾判别的精度。上述方法需要建立在猪只个体分割及其轮廓提取均正确的基础上,而在光照不均、猪只严重密集等复杂情况下很难保证轮廓提取的精度。不同于上述方法,文献<sup>[10-11]</sup>利用快速区域卷积神经网络(Fast region-based convolutional neural networks, Fast R-CNN)模型直接训练猪只头部检测器,取得了较高的检测精度。但这种方法无法处理由于遮挡、强光等因素导致头部不可见的情况,并且该方法对头部与猪只整体分开进行检测,需要对头部和其所属猪只进行关联,但文中对此考虑较为简单,无法处理复杂情况。

鉴于猪只头部和尾部在判断猪只行为中的重要作用,并且二者可以基本确定猪只身体方向,本文选取猪只头部和尾部为两类关键部位,从群养猪只的监控视频中自动获取准确的猪只及其头尾位置信息,提出一种猪只检测及头尾辨别方法,以提高群养猪只自动行为识别及监测水平。

## 1 材料与方法

### 1.1 视频采集

本文的实验图像均采集于南京农业大学江浦农

场生猪养殖基地。所选取的实验猪舍面积约为  $3.8\text{ m} \times 3.2\text{ m}$ ,猪舍中央安装的铁栏杆将整个猪舍分割为两个独立的小型猪舍,每个小型猪舍中饲养 4 头保育期猪只。监控摄像机安装于实验猪舍中心位置的天花板处,近似位于铁栏杆的正上方,高度约  $2.5\text{ m}$ 。监控摄像机以俯视的角度进行拍摄。选用镜头焦距为  $4\text{ mm}$  的海康威视 DS-2CD3335-I 型红外摄像机。由于受猪舍高度及摄像机镜头焦距限制,摄像机所能拍摄到的监控画面并不能完整覆盖整个猪舍,因此靠近画面上侧与左侧猪舍边缘处的猪只可能会出现整体或部分身体超出摄像机成像范围的情况。

实验期间摄像机连续采集 5 d 猪只活动视频。考虑到夜晚时段的猪只多处于睡眠状态,因此仅选取每日 08:00—17:00 间猪只较活跃时段的视频作为实验素材。摄像机工作过程中,监控视频存储于硬盘录像机中,帧率为  $12\text{ f/s}$ 。视频采集完成后,以 20 s 间隔从视频中抽取 540 幅图像作为图像数据集,图像分辨率为  $2\,048\text{ 像素} \times 1\,536\text{ 像素}$ 。从 540 幅图像中随机选取 300 幅图像作为训练集,其余 240 幅图像作为测试集。本文仅关注图像中头尾均完整可见的猪只目标。在训练集和测试集中完整猪只分别为 1 956 头和 1 584 头。在图像中用矩形框人工标注出所有完整猪只个体的身体、头部、尾部,便于后续模型训练和测试。由于所有的实验图像数据均采集于自然环境下的群养圈舍,图像中呈现的许多特点反映了实际生产环境的特点,例如不均匀的光照、图像中央的栏杆、多变的猪只姿态,这些因素提高了猪只检测和头尾辨别的难度,因此该图像测试集可用于客观评价本文方法在实际生产应用中的性能。

### 1.2 研究方法

本文提出了一种在群养猪只的监控图像中检测猪只目标并辨别其头尾位置的方法。该方法借鉴了计算机视觉中研究人体检测及姿态估计问题的一种思路,即引入基于部位的模型(Part-based model)<sup>[18-19]</sup>,将猪只看作由关键部位按一定空间关系组合成的目标。本文方法流程如图 1 所示。

首先,在输入图像中检测猪只整体、猪只头部、猪只尾部 3 种目标位置,作为进一步分析的基础。综合考虑检测精度及运行速度,选择基于深度神经网络的目标检测模型 YOLO v3(You only look once v3)<sup>[20]</sup>作为目标检测器,检测结果在图像中用矩形框表示。接下来,逐个提取猪只整体的检测矩形框判断其内部的部位情况。统计矩形框中心处于整体猪只矩形框内部的头部和尾部数目,分别用  $N_{\text{head}}$  和  $N_{\text{tail}}$  表示。

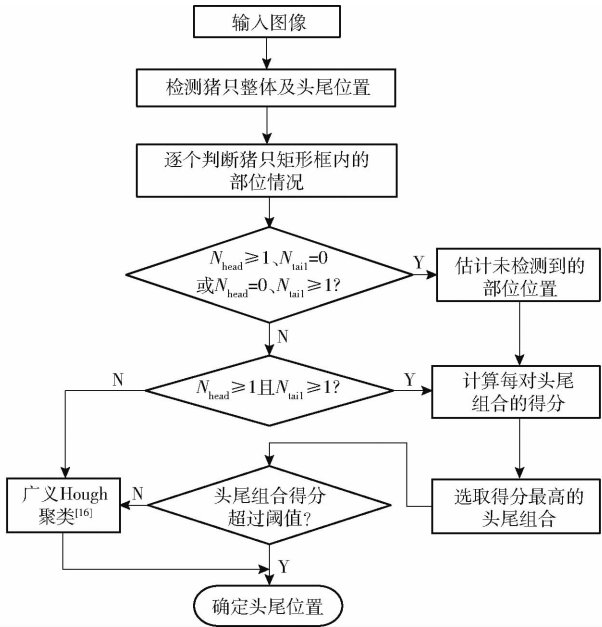


图 1 本文方法流程图

Fig. 1 Flow chart of proposed method

根据  $N_{head}$  和  $N_{tail}$  的取值,分为以下 3 种情况:

(1)  $N_{head} \geq 1$  且  $N_{tail} \geq 1$ ,即猪只整体矩形框内部检测到头尾两种部位,但可能出现多种头尾组合方法。这种情况下,利用预先建立好的图结构模型(Pictorial structure models)<sup>[21]</sup>对每种可能的头尾组合计算其组合得分,最后选取得分最高且超过阈值  $t_{score}$  的头尾组合作为当前猪只的头尾位置。

(2)  $N_{head} \geq 1$ 、 $N_{tail} = 0$  或  $N_{head} = 0$ 、 $N_{tail} \geq 1$ ,即猪只整体矩形框内部仅检测到头部或尾部一种部位。这种情况下需要根据当前图像的特征推理出缺失部位的位置。获得推理的部位位置后,接下来的处理方式与情况(1)相同。

(3)  $N_{head} = 0$  且  $N_{tail} = 0$ ,即猪只整体矩形框内未检测到任何部位。此时根据文献[16]中基于广义 Hough 聚类的方法作为一种弥补措施,辨别出头部和尾部的的位置。同时,当前两种情况下无得分超过阈值的部位组合时,也利用该方法确定头尾的位置。

通过上述 3 种处理方式,可以对检测到的每头猪只目标辨别出其头部和尾部的的位置。

1.3 基于 YOLO v3 的猪只整体及头尾检测

YOLO<sup>[22]</sup>是基于卷积神经网络的通用目标检测模型,该模型的特点是在进行检测时使用回归思路预测所检测目标的矩形框,因此与其他需要先进行选择搜索提取大量候选区域的检测方法相比,能够大大提高计算效率。YOLO v3 是 YOLO 系列模型的最新版本,借鉴了残差网络形成更深的网络层次,还融合特征金字塔网络改善了目标检测的性能,尤其适于小目标的检测。因此,从检测精度和效率两

方面综合考虑,本文选用 YOLO v3 模型检测猪只整体和头尾部位。

一般深度学习模型需要大量训练样本,而本文为了适应有限的训练样本,在模型训练时使用迁移学习的策略,引入在 VOC 数据集上预训练好的模型。在预训练模型的基础上,使用本文训练集中的标注数据进行模型参数微调,最终获得训练好的检测模型。该模型具有从输入图像中检测猪只整体及其头部、尾部 3 类目标的矩形框坐标及置信度的功能。

图 2 为测试集中猪只整体及部位检测结果示例。图中猪只整体检测结果用红色矩形框标示,头部和尾部检测结果分别用蓝色和绿色矩形框标示。为了方便叙述,对图像中的猪只标记了序号。由于 YOLO v3 模型的检测精度较高,图像中的大部分猪只及其部位都能被准确检测到。但在某些特殊情况,如遮挡、快速运动模糊等因素的影响下,仍然存在一些部位漏检的情况。如图 2 中 2 号猪只的尾部由于被 1 号猪只遮挡而未能检测到。另外,即使头尾能够被准确检测,但如何实现头尾与猪只整体间的正确关联也是一个重要的问题。如 3~6 号猪只的检测矩形框内除了含有相应序号猪只的部位外,还包含了其他猪只的部位。因此,还需要对部位检测结果进行进一步分析,才能够最终确定猪只头尾位置及其正确关联的整体猪只。



图 2 猪只整体及部位检测结果示例

Fig. 2 Example of detection results for pig and their parts based on YOLO v3

1.4 基于图结构模型的猪只部位组合

若从俯视角度观察猪只,会发现其头部和尾部具有一定的几何约束关系。这种约束可以作为一种先验知识,用于衡量头部和尾部的组合方式属于同一头猪的可信度。本文引入图结构模型来实现部位组合的功能。图结构模型是用一组部位和它们之间的联系来描述整体目标内部部位组合关系的模型,广

泛用于人脸、人体等可用不同部位的组合来描述的对象<sup>[23-24]</sup>。图结构模型中的部位代表了整体目标的局部视觉特征,部位的选取通常符合目标的语义分割方式,如将人体划分为头部、躯干、四肢等部位。而部位间用类似弹簧的模型描述部位间的形变特性,将部位的空间组合特征定义为一种形变配置(Deformable configuration)。这样,图结构模型就可以灵活表示各种形变配置情况下的整体目标。

本文所构建的猪只图结构模型如图 3 所示。整个模型的结构可看作一个简单的无向图,由 3 种类型的结点组成:整体猪只结点  $V_p$ 、头部结点  $V_h$ 、尾部结点  $V_t$ 。每个结点的配置信息由其图像坐标决定。在已知 3 种结点的图像坐标后,可按图 3 的方式构造同样结构的无向图,并定义得分函数  $\psi(V_p, V_h, V_t)$  计算当前图结构与预先建立的图结构模型的匹配程度。匹配得分为头尾部位的局部外观特征得分与部位之间几何约束得分的加权叠加。

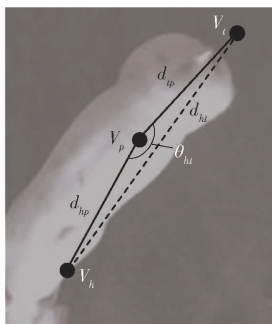


图 3 猪只图结构模型示意图

Fig.3 Pictorial structure model for pigs

$$\psi(V_p, V_h, V_t) = w_a \varphi_a(V_h, V_t) + w_g \varphi_g(V_p, V_h, V_t) \quad (1)$$

其中  $\varphi_a(V_h, V_t) = (c(V_h) + c(V_t))/2 \quad (2)$

$$\varphi_g(V_p, V_h, V_t) = w_{hp} N(d_{hp\_norm}; \mu_{hp}, \sigma_{hp}) + w_{tp} N(d_{tp\_norm}; \mu_{tp}, \sigma_{tp}) + w_{\theta} N(\theta_{ht}; \mu_{\theta}, \sigma_{\theta}) \quad (3)$$

$$\begin{cases} d_{hp\_norm} = d_{hp}/d_{ht} \\ d_{tp\_norm} = d_{tp}/d_{ht} \end{cases} \quad (4)$$

式中  $\varphi_a(\cdot)$ ——部位外观特征得分函数  
 $\varphi_g(\cdot)$ ——部位间几何约束得分函数  
 $w_a, w_g$ ——部位外观特征得分、部位间几何约束得分权重  
 $c(\cdot)$ ——部位的检测置信度,由 YOLO v3 模型输出,取值为 $[0, 1]$

$N(\cdot)$ ——高斯概率密度函数

$d_{hp\_norm}$ ——头部与猪只中心点间的归一化距离

$d_{tp\_norm}$ ——尾部与猪只中心点间的归一化距离

$\theta_{ht}$ ——头、尾部与猪只中心连线形成的夹角

$\mu_{hp}, \mu_{tp}, \mu_{\theta}$ ——高斯分布均值

$\sigma_{hp}, \sigma_{tp}, \sigma_{\theta}$ ——高斯分布的标准差

$w_{hp}, w_{tp}, w_{\theta}$ ——3 项得分对应的权重

$d_{hp}$ ——头部与猪只中心点之间的欧氏距离

$d_{tp}$ ——尾部与猪只中心点之间的欧氏距离

$d_{ht}$ ——头部与尾部间的欧氏距离

头尾部位的外观特征得分  $\varphi_a$  表示头尾结点检测置信度的平均值。部位间的几何约束得分为 3 项得分的加权求和:头部与猪只中心点之间的距离约束得分;尾部与猪只中心点之间的距离约束得分;头部、尾部与猪只中心点之间的方向约束得分。为了灵活描述图结构模型中部位间的空间形变约束,几何约束得分中的 3 个得分项均采用高斯概率密度函数计算部位组合的得分。高斯分布的参数由猪只先验结构知识与训练集中参数的统计规律综合确定。为了保证高斯概率密度函数的参数不受猪只尺寸的影响,将猪只头尾距离进行归一化。

当确定猪只检测矩形框内的头尾组合方式后,利用式(1)可计算当前组合得分,得分反映了部位配置与预先建立的图结构模型的匹配程度。当整体猪只检测矩形框内包含多种可能的头尾组合时,可通过不同组合方式的得分确定正确关联的头尾位置。以图 2 中的 4 号猪只为例,其检测矩形框中包含 2 个头部检测结果,因此可形成两种头尾组合方式,如图 4 所示。利用式(1)进行计算,图 4a 的头尾组合得分为 0.875 1,而图 4b 的头尾组合得分为 0.374 6,因此判定图 4a 的组合方式为最终确定的头尾位置。

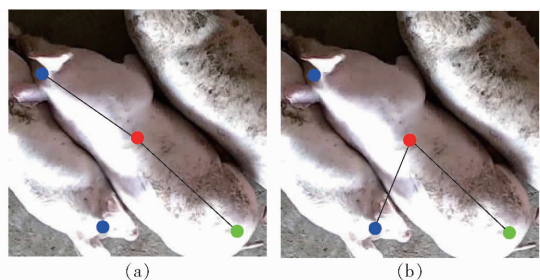


图 4 同一检测矩形框内的 2 种头尾组合示例

Fig.4 Examples for two combinations of heads and tails in one detection bounding box of a pig

## 1.5 基于阈值分割与椭圆拟合的部位推理

1.4 节中的部位组合方法要求猪只整体检测矩形框内同时存在头部和尾部两种部位。当其中一种类型的部位缺失时,需要推理出缺失部位的位置才能构建出完整的图结构。因此,本文提出了一种利用猪只身体的灰度特性与形状特性推理缺失部位位置的方法,方法流程如图 5 所示。

首先,提取猪只检测矩形框内的图像,使用最大熵阈值分割法<sup>[25]</sup>对该图像进行二值化操作。最大熵阈值分割法可以根据图像中像素灰度的统计特性



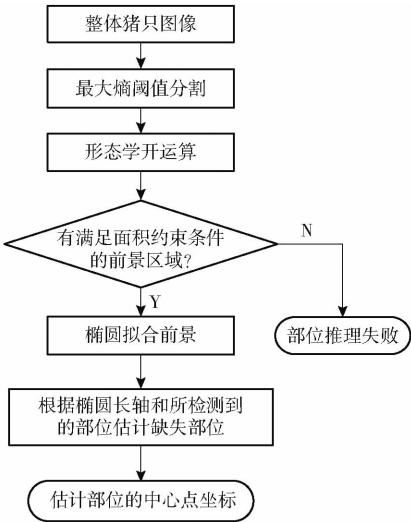


图 5 缺失部位推理流程图

Fig. 5 Flow chart of missing part inference

自动确定阈值,无需设定参数即可分割出具有较亮灰度的前景像素。然后,进行形态学开运算操作,目的是去除前景中小面积的连通分量区域,且尽可能将粘连区域分割。接下来,选出图像中面积最大且面积超过一定阈值的连通分量区域,即满足

$$A_{\max} \geq \alpha W_{\text{img}} H_{\text{img}} \tag{5}$$

式中  $A_{\max}$ ——连通分量区域的最大面积

$\alpha$ ——最大面积阈值

$W_{\text{img}}$ ——整体猪只图像的宽度

$H_{\text{img}}$ ——整体猪只图像的高度

若存在满足面积约束条件的连通分量,则对该区域进行椭圆拟合,获取拟合椭圆的各项参数。根据猪只的身体形状特点,理想情况下猪只的头部和尾部应该处于拟合椭圆长轴的两端。因此,在已知椭圆中心点坐标  $(x_0, y_0)$ 、长轴长度  $a$ 、椭圆倾斜角  $\varphi$  的条件下,根据

$$\begin{cases} x_1 = x_0 + \frac{a}{2} \beta \cos \theta_e \\ y_1 = y_0 + \frac{a}{2} \beta \sin \theta_e \end{cases} \tag{6}$$

$$\begin{cases} x_2 = x_0 - \frac{a}{2} \beta \cos \theta_e \\ y_2 = y_0 - \frac{a}{2} \beta \sin \theta_e \end{cases} \tag{7}$$

式中  $(x_1, y_1)$ ——推断头部中心点坐标

$(x_2, y_2)$ ——推断尾部中心点坐标

$\beta$ ——头尾位置分割椭圆长轴的比例,为常数

$\theta_e$ ——椭圆长轴与水平方向的夹角

推断出两个备选部位中心点  $(x_1, y_1)$  和  $(x_2, y_2)$ 。最后,选取两个备选部位点中距已知所检测到的部位较远的点为推理结果。部位推理结果与所检测到的部位一同构成一种头尾组合,当存在多种头尾组合

时可利用 1.3 节的方法选取最优的组合方式作为头尾辨别的结果。

以图 2 中的 2 号猪只为例,图 6 展示了尾部推理过程。图中推理出的部位中心点用虚线框标记,可以看到被遮挡的尾部能够被正确推理。

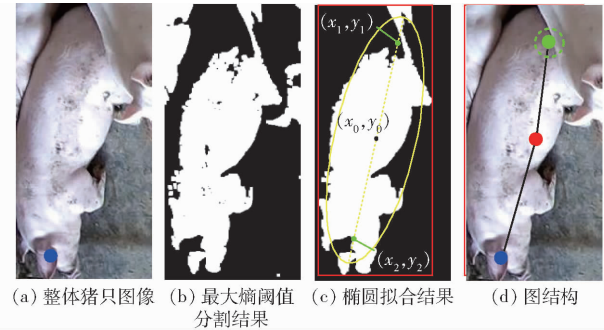


图 6 部位推理过程示例

Fig. 6 Example for process of part inference

经过上述方法的处理,图 2 中的猪只均能正确分辨其头尾位置,结果如图 7 所示。



图 7 对图 2 中所检测到猪只的头尾辨别结果

Fig. 7 Head/tail identification results for Fig. 2

## 2 实验结果与分析

### 2.1 实验配置与评价指标

实验所用的硬件处理平台为台式计算机,配置为 Intel i7 - 8700K 3.7 GHz CPU, GPU 为 GTX1080Ti,带有 11 GB 独立显存和 16 GB 内存。实验中 YOLO v3 模型的学习和测试部分在开源的 Darknet 框架下使用 GPU 运行,对部位检测结果的处理分析,如计算头尾组合得分、部位推理等操作则在 Matlab 平台上运行。实验中的关键参数如表 1 所示。

使用精确率 (Precision) 和召回率 (Recall) 两个指标衡量猪只整体及部位的定位精度。需要注意的是,对猪只整体来说,若检测矩形框与预先标定的真实矩形框的交并比 (Intersection over union) 大于 0.5,则定义为检测正确;而对于检测或推理获得的头尾位置,其中心点处于预先标记的真实头尾矩形框范围内,则认为部位定位正确。另一评价指标为

表 1 实验中的关键参数

Tab. 1 Settings of important parameters in experiments

| 参数                         | 数值   | 含意                |
|----------------------------|------|-------------------|
| $t_{yolo}$                 | 0.2  | YOLO v3 检测结果阈值    |
| $w_a$                      | 0.3  | 局部外观得分权值          |
| $w_g$                      | 0.7  | 部位间几何约束得分权值       |
| $\mu_{hp}, \mu_{tp}$       | 0.5  | 距离约束高斯分布均值        |
| $\mu_{\theta}$             | 180  | 方向约束高斯分布均值        |
| $\sigma_{hp}, \sigma_{tp}$ | 0.15 | 距离约束高斯分布标准差       |
| $\sigma_{\theta}$          | 18   | 方向约束高斯分布标准差       |
| $w_{hp}, w_{tp}$           | 0.3  | 距离约束得分权值          |
| $w_{\theta}$               | 0.4  | 方向约束得分权值          |
| $\alpha$                   | 0.5  | 部位推理时最大前景面积阈值     |
| $\beta$                    | 0.7  | 部位推理时部位点分割椭圆长轴的比例 |
| $t_{score}$                | 0.65 | 部位组合得分阈值          |

猪只头尾辨别的精确率,即计算头尾与猪只整体正确关联且头尾位置均准确的猪只占有检测到猪只的比例。

### 2.2 猪只头尾辨别结果

为了验证本文方法的有效性,本文在测试集上

表 2 部位组合及推理前后猪只整体和部位的检测结果

Tab. 2 Detection results for pigs and their parts before and after part assemble and inference

%

| 检测方法            | 整体猪只  |       | 头部    |       | 尾部    |       |
|-----------------|-------|-------|-------|-------|-------|-------|
|                 | 精确率   | 召回率   | 精确率   | 召回率   | 精确率   | 召回率   |
| YOLO v3 模型      | 99.30 | 98.42 | 94.97 | 96.46 | 94.22 | 94.70 |
| YOLO v3 + 图结构模型 |       |       | 98.23 | 98.23 | 97.85 | 97.85 |

本文方法、文献[10]方法、文献[16]方法的猪只头尾辨别精确率分别为 96.22%、91.53%、79.41%。需要注意文献[10]方法仅实现了头部与整体的关联,且使用的检测模型为 Fast R-CNN。为了便于比较,在实验中将尾部也用同样的方法与整体猪只相关联,且将检测模型替换为 YOLO v3,这样更能在同等条件下证明本文所提出的图结构模型与部位推理方法的作用及意义。另外,实验中使用文献[16]的方法时,利用了本文的部位推理方法获取拟合椭圆,然后将椭圆长轴两端的前景区域作为猪只身体两端的轮廓,用于计算广义 Hough 聚类的结果。这里与原文的轮廓提取方法有细微差别。本文方法相对于另外两种对比方法在头尾辨别精确率上具有明显优势。相对于文献[10]中的方法,本文方法的优势在于增加了构建图结构模型计算头尾组合得分及部位推理两部分,因此可以处理猪只整体矩形框内存在多种头尾组合方式或部位漏检等特殊情况,而文献[16]的方法需要在头尾轮廓准确提取的基础上实现。本文实验的测试图像拍摄于光照不均匀、背景干扰复杂的圈舍环境下,使用阈值分割方

进行了实验。表 2 为直接利用 YOLO v3 模型检测猪只整体及部位的结果和在此基础上进一步利用图结构模型进行部位组合和部位推理后的结果。从表 2 可以看出,YOLO v3 模型在检测猪只整体方面可以获得较理想的结果。但由于头、尾部位的图像尺寸较小,使得图像中包含的视觉特征不足,头部和尾部检测的精确率和召回率均低于整体猪只的检测精度。同时也注意到,尾部检测的精确率和召回率略低于头部,这也符合图像的特点,即尾部图像相对于头部图像更缺乏明显的视觉特征,且相对于头部更容易因拥挤被遮挡,导致尾部检测的漏检增多。表 2 说明经过部位组合和部位推理等操作后,头部和尾部检测的精确率和召回率都有明显提升。一方面是因为成功推理出了部分缺失部位,另一方面是由于一些部位检测虚警(False alarms),未能与整体猪只相关联而被去除。由于部位组合后每头猪只整体都能确认一对头尾组合,因此漏检部位数与虚警部位数是相同的,这时精确率与召回率相等。

法很难获得完全精确的猪只前景区域,因此依靠前景轮廓判断头尾也很难获得准确结果。图 8 为测试集中猪只头尾辨别结果示例。图 8 中左图为本文方法辨别头尾的结果,右图为文献[16]方法的辨别结果,其中蓝色和绿色框分别为头部和尾部的辨别结果。图 8a 和图 8c 为检测部位缺失时的情况,图 8b 和图 8d 为整体猪只矩形框内存在多种部位组合的情况。从图 8 中可直观感受到本文方法相对于其他两种方法的优势。需要注意的是,本文在猪只整体矩形框内没有检测到任何部位时,使用文献[16]中的方法作为补偿方案。由于这种情况非常少见,因此少量的精度损失是可接受的。

本文所提出的猪只头尾辨别方法是建立在对猪只整体及其头尾准确定位的基础上,即确保 YOLO v3 模型的检测精度。作为一个深度神经网络模型,决定 YOLO v3 模型检测性能的重要因素为训练样本的质量。因此,若要将本文方法推广应用到其他俯视拍摄的视频监控场景中,只需将该场景下的猪只样本图像进行标记,并用其对 YOLO v3 模型进行参数微调即可。图 9 为另一个场景中,对 YOLO v3

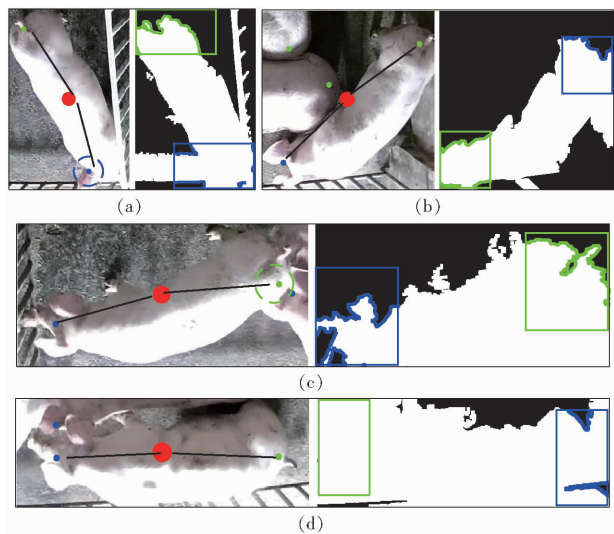


图8 头尾辨别结果示例

Fig. 8 Examples of pig head/tail identification

模型进行参数微调并结合图结构模型对猪只头尾辨别的结果。从图9中可以看出,本文方法依然能正确辨别出猪只头尾位置,证明了本文方法具有良好

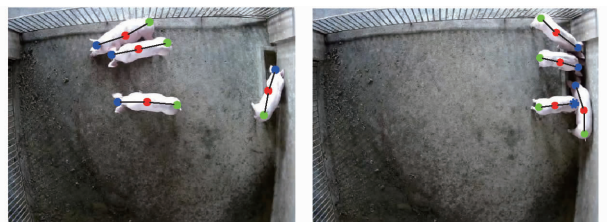


图9 另一场景下的猪只头尾辨别结果示例

Fig. 9 Examples of pig head/tail identification results in another scene

的泛化性能。

### 3 讨论

虽然本文方法取得了可靠的猪只头尾辨别效果,但依然存在一定的缺点。首先,本文方法必须建立在猪只整体正确检测的基础上,且最终头尾辨别的正确率高度依赖头尾检测的精度。若部位漏检率较高,或存在大量的虚警目标,会降低头尾辨别的精度。其次,本文方法限定一个猪只整体检测矩形框内仅存在一种正确的头尾组合,若猪只过于密集,导致一个矩形框内存在两个完整的猪只身体时,本文无法辨别两只猪各自的头尾组合。另外,本文所用的部位推理方法依然较为基础,虽然目前作为一种弥补手段暂时可行,但若希望继续提高算法精度,仍需要进行更深入的考虑。

### 4 结束语

提出了一种基于YOLO v3与图结构模型的群养猪只头尾辨别方法。在实际圈舍群养环境下的监控图像集上进行了实验,结果表明,本文方法比YOLO v3模型在头尾定位的精确率和召回率上均有明显提高,本文方法的头尾辨别精确率可达96.22%,优于文献[10]和文献[16]的方法。因此,本文方法可有效实现群养猪只的准确检测及头尾辨别,受光照、猪只头数、姿态等干扰因素的影响较小。

### 参考文献

- [1] 王德福,黄会男,张洪建,等.生猪养殖设施工程技术研究现状与发展分析[J/OL].农业机械学报,2018,49(11):1-14. WANG Defu, HUANG Huinan, ZHANG Hongjian, et al. Analysis of research status and development on engineering technology of swine farming facilities[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(11): 1-14. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20181101&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20181101&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2018.11.001. (in Chinese)
- [2] 沈明霞,刘龙申,闫丽,等.畜禽养殖个体信息监测技术研究进展[J/OL].农业机械学报,2014,45(10):245-251. SHEN Mingxia, LIU Longshen, YAN Li, et al. Review of monitoring technology for animal individual in animal husbandry[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 245-251. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20141038&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20141038&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2014.10.038. (in Chinese)
- [3] 孙红,李松,李民赞,等.农业信息成像感知与深度学习应用研究进展[J/OL].农业机械学报,2020,51(5):1-17. SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 1-17. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?file\\_no=20200501&flag=1](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20200501&flag=1). DOI:10.6041/j.issn.1000-1298.2020.05.001. (in Chinese)
- [4] OTT S, MOONS C, KASHIHA M, et al. Automated video analysis of pig activity at pen level highly correlates to human observations of behavioural activities[J]. Livestock Science, 2014, 160: 132-137.
- [5] NASIRAHMADI A, EDWARDS S A, MATHESON S M, et al. Using automated image analysis in pig behavioural research: assessment of the influence of enrichment substrate provision on lying behaviour[J]. Applied Animal Behaviour Science, 2017, 196: 30-35.
- [6] 李丹,张凯锋,李行健,等.基于Mask R-CNN的猪只爬跨行为识别[J/OL].农业机械学报,2019,50(增刊):261-266, 275. LI Dan, ZHANG Kaifeng, LI Xingjian, et al. Mounting behavior recognition for pigs based on Mask R-CNN[J/OL].

- Transactions of the Chinese Society for Agricultural Machinery, 2019, 50 (Supp. ): 261 – 266, 275. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=2019s041&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2019s041&journal_id=jcsam). DOI: 6041/j. issn. 1000-1298. 2019. S0. 041. (in Chinese)
- [7] ZHU W, GUO Y, JIAO P, et al. Recognition and drinking behaviour analysis of individual pigs based on machine vision[J]. *Livestock Science*, 2017, 205: 129 – 136.
- [8] KASHIHA M, BAHR C, HAREDasht S, et al. The automatic monitoring of pigs water use by cameras[J]. *Computers and Electronics in Agriculture*, 2013, 90: 164 – 169.
- [9] NASIRAHMADI A, HENSEL O, EDWARDS S, et al. Automatic detection of mounting behaviours among pigs using image analysis[J]. *Computers and Electronics in Agriculture*, 2016, 124: 295 – 302.
- [10] YANG Q, XIAO D, LIN S. Feeding behavior recognition for group-housed pigs with the Faster R – CNN[J]. *Computers and Electronics in Agriculture*, 2018, 155: 453 – 460.
- [11] 杨秋妹, 肖德琴, 张根兴. 猪只饮水行为机器视觉自动识别[J/OL]. *农业机械学报*, 2018, 49(6): 232 – 238. YANG Qiumei, XIAO Deqin, ZHANG Genxing. Automatic pig drinking behavior recognition with machine vision [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(6): 232 – 238. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20180627&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180627&journal_id=jcsam). DOI: 10. 6041/j. issn. 1000-1298. 2018. 06. 027. (in Chinese)
- [12] GUO Y, ZHU W, JIAO P, et al. Multi-object extraction from topview group-housed pig images based on adaptive partitioning and multilevel thresholding segmentation[J]. *Biosystems Engineering*, 2015, 135: 54 – 60.
- [13] BRUNGER J, TRAULSEN I, KOCH R. Model-based detection of pigs in images under sub-optimal conditions[J]. *Computers and Electronics in Agriculture*, 2018, 152: 59 – 63.
- [14] LI B, LIU L, SHEN M, et al. Group-housed pig detection in video surveillance of overhead views using multi-feature template matching[J]. *Biosystems Engineering*, 2019, 181: 28 – 39.
- [15] 沈明霞, 太猛, CEDRIC Okinda, 等. 基于深层卷积神经网络的初生仔猪目标实时检测方法[J/OL]. *农业机械学报*, 2019, 50(8): 270 – 279. SHEN Mingxia, TAI Meng, CEDRIC Okinda, et al. Real-time detection method of newborn piglets based on deep convolution neural network[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(8): 270 – 279. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20190830&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190830&journal_id=jcsam). DOI: 10. 6041/j. issn. 1000-1298. 2019. 08. 030. (in Chinese)
- [16] 杨心, 朱伟兴. 基于广义 Hough 聚类的粘连猪头尾识别[J]. *江苏农业科学*, 2018, 46(9): 230 – 235. YANG Xin, ZHU Weixing. A recognition method for head and tail of adhesive pigs based on generalized Hough clustering[J]. *Jiangsu Agricultural Sciences*, 2018, 46(9): 230 – 235. (in Chinese)
- [17] 高云, 郁厚安, 雷明刚, 等. 基于头尾定位的群养猪运动轨迹追踪[J]. *农业工程学报*, 2017, 33(2): 220 – 226. GAO Yun, YU Houan, LEI Minggang, et al. Trajectory tracking for group housed pigs based on locations of head/tail [J]. *Transactions of the CSAE*, 2017, 33(2): 220 – 226. (in Chinese)
- [18] SUN M, SAVARESE S. Articulated part-based model for joint object detection and pose estimation [C] // *International Conference on Computer Vision and Pattern Recognition*, 2011: 723 – 730.
- [19] LI B, CHEN Y, WANG F Y. Pedestrian detection based on clustered poselet models and hierarchical AND – OR grammar [J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(4): 1435 – 1444.
- [20] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2018 – 04 – 08]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [21] FELZENSZWALB P. Pictorial structure for object recognition[J]. *International Journal on Computer Vision*, 2005, 61(1): 55 – 79.
- [22] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779 – 799.
- [23] LÓPEZ-QUINTERO M, MARÍN-JIMÉNEZ M, MUÑOZ-SALINAS R, et al. Stereo pictorial structure for 2D articulated human pose estimation[J]. *Machine Vision and Applications*, 2016, 27(2): 157 – 174.
- [24] TAN X, SONG F, ZHOU Z, et al. Enhanced pictorial structures for precise eye localization under uncontrolled conditions [C] // *International Conference on Computer Vision and Pattern Recognition*, 2009: 1621 – 1628.
- [25] KUMAR S, PANT M, RAY A K. 2D maximum entropy method for image thresholding converge with differential evolution[J]. *Advances in Mechanical Engineering and Its Applications*, 2012, 2(3): 189 – 192.