

doi:10.6041/j.issn.1000-1298.2021.09.020

基于多尺度感知的高密度猪只计数网络研究

高云^{1,2} 李静¹ 余梅³ 李小平³ 余慧祥¹ 谭忠¹

(1. 华中农业大学工学院, 武汉 430070; 2. 生猪健康养殖协同创新中心, 武汉 430070;
3. 华中农业大学动物科技学院, 武汉 430070)

摘要: 猪只盘点是生猪规模化养殖和管理中的重要环节,人工计数方法费时、费力,在大数据量的猪只盘点中容易出错。本文使用多尺度感知网络对高密度猪群图像中的猪只进行计数。通过对人群计数网络 CSRNet 的改进,得到猪只计数网络(Pig counting net, PCN),PCN 采用 VGG16 作为前端网络提取特征,中间层采用空间金字塔(Spatial pyramid)结构对图像中的多尺度信息进行提取与融合,后端网络采用改进的膨胀卷积网络。PCN 增加了多尺度感知结构、扩大了后端网络感受野,通过感知多尺度特征得到预测密度图,预测密度图反映了猪只空间分布,通过对密度图积分实现了猪只数量的估计。结果表明,在平均猪只数为 40.71 的测试集图像上,PCN 的计数准确率优于人群计数网络 MCNN、CSRNet 和改进 Counting CNN 的猪只计数网络,MAE 和 RMSE 分别为 1.74 和 2.28,表现出较高的准确性和鲁棒性;单幅图像平均识别时间为 0.108 s,满足实时处理要求。

关键词: 猪只盘点; 高密度计数; 空间金字塔; 多尺度感知; 深度学习

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2021)09-0172-07 OSID: 

High-density Pig Counting Net Based on Multi-scale Aware

GAO Yun^{1,2} LI Jing¹ YU Mei³ LI Xiaoping³ YU Huixiang¹ TAN Zhong¹

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China
2. The Cooperative Innovation Center for Sustainable Pig Production, Wuhan 430070, China
3. College of Animal Science and Technology, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: Pig inventory is an important part of large-scale breeding and management of live pigs. Manual counting methods are more time-consuming and laborious, especially in pig inventory with large amounts of data. How to count high-density pig herd images with machine vision is still a difficult problem to be solved urgently. A multi-scale aware counting network was used to count pigs in high-density pig herd images. Based on the crowd counting network CSRNet, the pig counting network of pig counting net (PCN) was proposed. VGG16 was used as the front-end network to extract features, the spatial pyramid structure was used, and this structure can extract and fusion multi-scale information in the image, the back-end network used an improved dilated convolutional network. PCN added a multi-scale aware structure, expanded the back-end network receptive field, and can obtain a predicted density map by sensing multi-scale features, the predicted density map reflected the spatial distribution of pigs, then by integrating the density map, the number of pigs can be accurately calculated. The results showed that on the test set image with an average number of pigs of 40.71, the accuracy of PCN was better than that of the crowd counting net MCNN, CSRNet and the pig counting net that modified Counting CNN, the mean absolute error (MAE) and the root mean square error (RMSE) were 1.74 and 2.28, respectively, the lower error showed that PCN had better accuracy and robustness. The average recognition time of a single image of the final model was 0.108 s, which met the real-time processing requirements of the algorithm. The method provided a research idea for the automatic inventory of high-density group raising pigs.

Key words: pig inventory; high-density counting; spatial pyramid; multi-scale aware; deep learning

收稿日期: 2020-08-29 修回日期: 2020-11-23
基金项目: 国家重点研发计划项目(2016YFD0500506)、中央高校自主创新基金项目(2662018JC003、2662018JC010、2662017JC028)和现代农业产业技术体系项目(CARS-35)
作者简介: 高云(1974—),女,副教授,博士,主要从事农业智能检测与控制研究,E-mail: angelclouder@mail.hzau.edu.cn

0 引言

猪只计数,俗称盘点,是集约猪场规模化管理中不可缺少的环节。养殖过程中猪只数量在不断发生变化,因此在管理过程中需要对猪只数量进行统计。常用的猪只计数方法往往依赖于人工,其计数结果容易产生误差,且工作量大、费时、费力。目前,规模化猪场的猪只数量越来越多,饲养密度也逐年提高,这对高密度猪只盘点提出了新的要求。采用自动化方法对高密度猪只进行计数有助于节约人工成本、提高管理效率。

目前,机器视觉方法由于非接触、低成本、实现容易等特点,已成为智能化、自动化猪只盘点的主要技术方法^[1]。传统的机器视觉猪只计数方法通过提取图像中的轮廓、颜色、形状等特征,将猪只与背景分割开,通过目标检测可以直接对图像中的猪只数量进行判断^[2-3],这种方法提取的特征较为单一,适用于猪只较少的简单场景。而实际猪场中采集到的猪群图像猪只密度较大、场景较为复杂,使用简单场景下的计数模型存在迁移困难的问题。近年来,深度学习方法在复杂场景中可以提取到更高级的特征,比机器视觉方法拥有更优异的性能,在生猪养殖领域得到了广泛的应用^[4-6]。基于深度学习的猪只计数方法不再依赖于前景分割和人工提取特征,TIAN 等^[7]采用深度学习方法改进了计数卷积神经网络,克服了猪只拥挤、光照不良的问题。

近年来,深度学习方法在处理高密度计数问题上显示了不可比拟的优势^[8-10],在人群计数研究中取得了良好的进展,可以有效预测高密度的人群图像^[11-12]。实际猪场中采集到的猪只数据集密度较大,但与人群数据集相比,其空间信息更加简单,轮廓也更加清晰,因此,将人群计数网络应用到猪只计数上是可行的。许多学者进行了人群计数方面的研究。目标检测^[13-15]和回归^[16]方法直接通过特征到数量的映射进行计数,可以对简单的人群场景进行数量识别。密度估计方法通过学习图像特征和相应的密度图之间的映射^[17-18],并对密度图进行积分得到计数结果,同时保留了计数目标的空间分布信息。基于深度学习的计数方法综合了密度估计方法,采用端到端的思想搭建卷积神经网络,通过网络学习将输入图像的特征映射到对应的密度图上,同样可以得到目标计数结果及其空间分布^[19-21]。研究证明,在场景复杂、密度较大的数据集上,基于深度学习的人群计数方法优于其他人群计数方法^[22]。而对于镜头透视带来的计数目标空间分布不均的问题,通过搭建网络、提取图像中的多尺度特征,可以

整合图像中的空间信息,进而得到更小的计数误差^[23]。

针对高密度猪只计数存在的问题,本文改进人群计数网络 CSRNet^[21],通过添加空间金字塔结构对图像中不同空间尺度的信息进行感知,建立猪只计数网络(Pig counting net, PCN),通过对网络进行不断训练与优化,得到端到端的猪只计数模型。

1 材料与方法

1.1 图像采集

于 2019 年 6 月在南方某集约化猪场进行数据采集,地点为断奶仔猪的赶猪通道,采用小米 8 智能手机拍摄,摄像头俯视拍摄角度为 30°左右。视频采集对象为 8 kg 左右的二元杂断奶仔猪,视频内的仔猪数量从 10 头至 180 头不等。试验共采集视频 13 段,视频长度在 10 ~ 36 s 之间,采用 mp4 格式存储,分辨率为 1 920 像素 × 1 080 像素。对采集到的视频以每秒 20 幅的速度进行图像截取,并剔除相似度过高、清晰度较差的图像。

1.2 数据集制作

为了保证样本的多样性,将得到的图像进行预处理,包括旋转、裁剪、打乱顺序。使用 Matlab R2016b 作为标定工具,编写图像标定程序,以坐标点的形式对每幅图像上的猪只进行标注,优先标定猪的脊背与头部交界处,以坐标(x,y)形式存储每头猪的位置。标签文件为 .mat 格式,标定点的个数即为图像中猪只的真实个数,标定点的坐标位置反映了猪只在图像中的真实空间分布信息^[24]。猪只计数数据集如表 1 所示。

表 1 数据集
Tab.1 Dataset

数据集	图像数量/ 幅	猪只总数	单幅图像最 大猪只数	平均每幅图 像猪只数
训练集	1 403	55 740	174	39. 73
验证集	350	13 776	168	39. 36
测试集	351	14 288	129	40. 71
合计	2 104	83 804	471	

试验标定了 2 104 幅图像,共有 83 804 头猪只,平均每幅图像 39. 83 头猪只,单幅图像中最大猪只数为 174。训练集、验证集与测试集图像数量比约为 4:1:1。

1.3 密度图生成

采用密度图计数方法,密度图反映猪只的空间分布信息,对密度图进行积分即可得到计数结果。为了生成密度图,需要对标签文件作进一步处理,对每个被标注的点进行高斯核卷积,将标注点平滑化,

生成高斯密度图,即真实密度图 (Ground truth)^[17]。以图像左上角为坐标原点,被标定的每头猪只坐标位置记为 1,其余为 0。考虑到透视失真带来的影响,使用 K 近邻算法,结合不同标注点横纵坐标之间的距离,得到每个个体与其临近个体的平均距离,并以此距离生成自适应高斯核,使用自适应高斯核对被标注点进行卷积。猪只密度图可以表示为标定点与自适应高斯核卷积所得到的结果,通过对高斯密度图进行积分可以得到图像中计数目标的数量。为了更直观地观测猪只的空间分布,使用热力图对高斯密度图进行可视化。原始图像为 RGB 三通道,而高斯密度图与网络预测的密度图则为单通道。

2 猪只计数模型的建立

以人群计数网络 CSRNet 为基础建立 PCN, CSRNet 在高度拥挤的人群计数场景中可以得到准确的计数结果和高质量的预测密度图。PCN 网络改进步骤如下:

- (1)增加 3 层空间金字塔结构。3 层空间金字塔结构对不同尺度的空间信息进行感知,得到尺度感知特征。
- (2)将尺度感知特征累加并与前端网络提取的特征进行融合。融合过程中得到多尺度特征,将前端网络与后端网络更好地连接。
- (3)改进 CSRNet 的膨胀卷积后端网络。扩大后端网络的感受野,提取深层信息,保持图像分辨率。

PCN 使用深层卷积神经网络,网络中间不包含全连接层,因此网络更加灵活,不需要固定网络输入图像的尺寸^[25],降低了网络的复杂性。网络具体结构如图 1 所示。

2.1 VGG16 前端

PCN 使用预训练的 VGG16 网络作为前端 (front-end) 网络,将 VGG16 的 13 层结构精简到了 10 层,并删除后面的全连接层。VGG16^[26]具有灵活性和传输学习能力强的优点,使用多层卷积配合 3×3 的卷积核,可以实现与大尺寸卷积核同样的感受野,且相较于大尺寸卷积核具有更少的参数量和更好的非线性变换能力^[27-30]。前端网络结构如图 1 所示。通过前端网络得到的输出特征图为原图尺寸的 1/8,便于 PCN 生成高质量的密度图。VGG16 前端网络得到的特征定义为 f_v ,作为中间层的基础输入特征。

2.2 多尺度感知中间层

PCN 使用空间金字塔结构搭建了多尺度感知中间层。前端网络提取到的特征在整幅图像上拥有

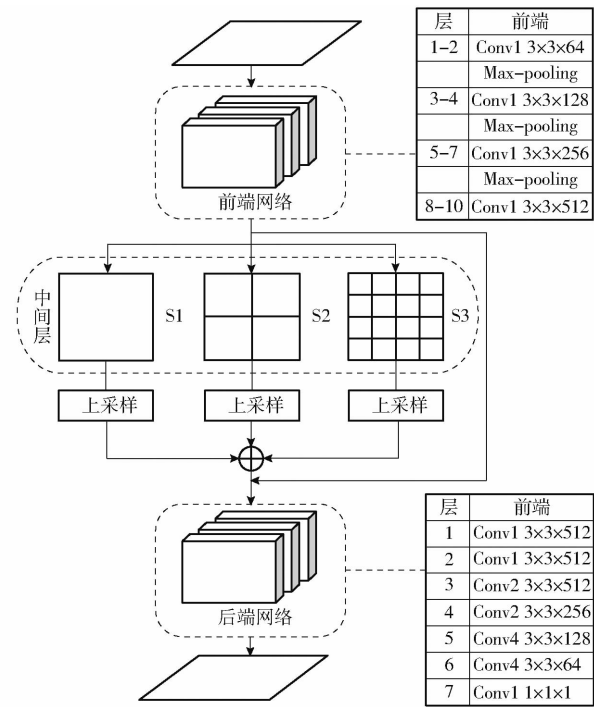


图 1 猪只计数网络结构图

Fig. 1 Structure diagram of pig counting net

相同的感受野,考虑到实际的猪只数据集中空间尺度连续变化问题,参照文献[30-32],PCN 将不同尺寸的空间金字塔池化层进行组合,对不同空间尺度的特征进行提取。猪只数据集与人群数据集相比,空间信息较为简单,计数目标轮廓更加清晰,因此 PCN 使用块尺寸更小、组合结构精简的空间金字塔,来提取猪只数据集中不同尺度的空间特征。

PCN 在前端网络后增加了 3 层空间金字塔池化层作为多尺度感知中间层,网络结构如图 1 所示, S1、S2、S3 表示 3 个不同块尺寸的空间金字塔池化层,3 个金字塔层分别将特征图划分为 1×1、2×2、4×4 的块,分别对这些块进行最大池化。在每个空间金字塔层后加一个卷积核为 1 的卷积层,以保证得到的尺度感知特征维度不变。使用双线性插值将 3 个尺度感知特征上采样到与 VGG 特征相同的大小,以统一不同尺度特征的尺寸,便于与 VGG 特征融合。通过空间金字塔结构对 VGG 特征进行卷积,得到 3 组不同的尺度感知特征 f_{si} ,将其累加并与 f_v 顺序连接,如图 1 所示,以得到空间信息更加全面的多尺度特征,多尺度特征 f 计算公式为

$$f = \left[f_v \mid \sum_{i=1}^3 f_{si} \right] \tag{1}$$

式中 i ——金字塔池化层编号

2.3 膨胀卷积网络后端

PCN 后端网络在原网络的基础上改进了膨胀率。猪只计数的目的是生成高质量的计数密度图便

于准确计数,而一般卷积操作后的池化层会降低图像分辨率,造成空间信息丢失,因此 PCN 的后端(back-end)网络使用膨胀卷积。膨胀卷积在融合多尺度特征的同时,可以保持网络输出密度图的分辨率^[33]。普通的卷积层通过使用更大的卷积核获得更大的视野,这也意味着网络参数数量的增加,膨胀卷积在扩大感受野且不增加参数数量的基础上可以保持图像分辨率,通过设置不同的膨胀率,膨胀卷积可以得到与大卷积核相同的感受野。

PCN 的后端网络共有 7 层,网络结构如图 1 所示,前 6 层均使用大小为 3×3 的卷积核,减少了后端网络的参数数量。PCN 改进了膨胀率为 2 的原后端网络,在多次试验后,前 6 层网络的膨胀率分别为 1、1、2、2、4、4,最后一层使用 1×1 卷积作为输出层。改进后的后端网络在猪只数据集上可以获得更大的感受野,避免了使用大卷积核扩大感受野带来的参数冗余。

3 猪只计数试验

3.1 试验平台搭建

试验在 64 位 Windows 10 操作系统的台式计算机上进行,运行内存为 16 GB,使用单 GPU 进行训练,GPU 型号为 RTX2060,显存为 6 GB。在此基础上,搭建了 GPU 版本的 pytorch 框架,使用 Python 编程实现了猪只计数网络的搭建。

3.2 模型训练参数

PCN 网络使用端到端的思想进行网络搭建和训练,模型的批尺寸(batchsize)设置为 1,使用随机梯度下降(SGD)作为优化算法,学习率设置为 1×10^{-6} ,动量(momentum)设置为 0.95。损失函数的计算公式为

$$L = \| D_{gt} - D_{est} \|^2 \tag{2}$$

式中 L ——损失函数

D_{gt} ——手动标记的真实密度图

D_{est} ——预测的密度图

在前向传播过程中,网络计算出预测密度图,并

将真实密度图与预测密度图逐像素相减,对各像素间的差值求平方后再相加,得到损失函数 L ,代表了预测密度图与真实密度图之间的差异。通过对网络的迭代,网络逐渐收敛, L 逐渐减小并趋于平稳,预测密度图越趋近于真实密度图,表示训练结果越准确。

3.3 猪只计数结果评价指标

深度学习的计数相关研究通常使用平均绝对误差(MAE)、均方根误差(RMSE)作为模型评价指标。MAE 为预测值与真实值的平均误差,表示算法的准确性,MAE 越小,算法准确度越高;RMSE 为预测值与真实值误差的分散程度,表示算法的稳定性, RMSE 越小,算法的鲁棒性越高^[34]。

3.4 猪只计数试验流程

多尺度感知猪只计数网络 PCN 的计数试验具体流程为:

(1)从集约化养殖猪场赶猪通道采集视频,从视频中截取图像,对图像进行预处理。

(2)使用 Matlab 程序对图像进行标定,生成每幅图像对应的标签文件,每头猪只的位置以坐标点的形式进行存储,标定好的数据按照 4:1:1 的比例划分训练集、验证集和测试集。

(3)编写 Python 程序,对标签文件进行自适应高斯卷积,生成图像的真实密度图。

(4)在 CSRNet 基础上改造猪只计数网络 PCN,并对网络结构进行调整与优化。

(5)对网络进行多次训练,选择合适的超参数;根据训练结果优化网络结构,获得最优的猪只计数网络模型。

3.5 试验结果与分析

3.5.1 密度图可视化结果

猪只预测密度图不仅可以用于计算猪只数量,还可以得到猪只分布情况。为了直观地显示网络计算得到的猪只空间分布,使用热力图对预测密度图进行可视化处理,并与真实密度图(Ground truth)对比,图 2 为不同情况下猪只计数网络密度图。

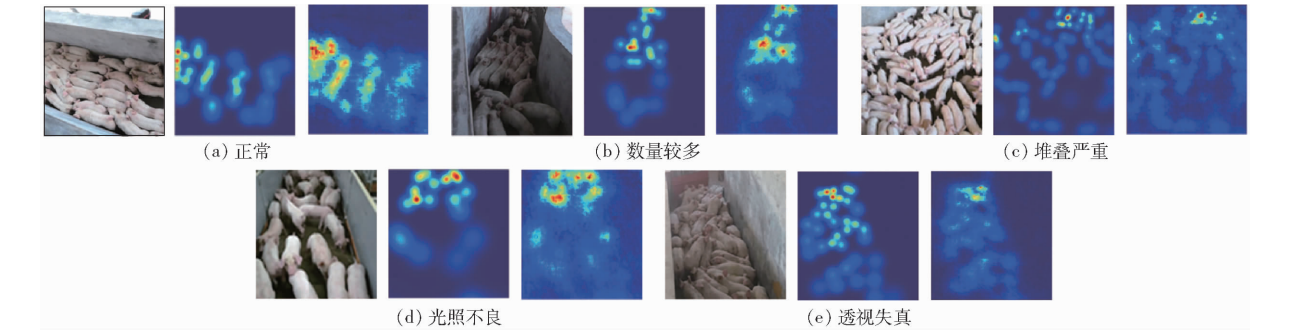


图 2 不同情况下的猪只计数网络密度图

Fig. 2 Density map of pig counting net on different condition

图 2 中从左到右依次为原图、真实密度图、预测密度图。其中预测密度图为 PCN 模型生成的猪只空间分布预测结果,尺寸为真实密度图的 1/8。图 2a~2e 原图中真实猪只数量为 29、24、81、22、41,预测猪只数量为 29.12、22.49、77.38、20.78、37.50。可以看出,PCN 模型对猪只的空间分布情况进行合理的预测。图 2a、2b、2d 预测值接近真实值,图 2c、2e 中猪只数量较多、堆叠严重,预测的猪只数量较为合理。

进一步分析光照不良、数量较多、透视失真、堆叠严重 4 种不良条件对模型造成的影响。将测试集按正常情况及 4 种不良条件进行分类,其中猪只数量大于 60 只的图像被归类为“数量较多”,测试集中有 16% 的图像同时存在 2 种及 2 种以上不良条件。使用 PCN 模型对不同条件下的猪只图像进行计数,测试结果统计如表 2 所示。

表 2 不同条件下猪只数量预测结果
Tab.2 Predicting results of number of pig sunder different conditions

条件	图像数量/幅	MAE	RMSE
正常	111	1.32	1.58
光照不良	55	1.47	1.95
数量较多	91	2.35	2.82
透视失真	86	1.79	2.16
堆叠严重	67	2.23	2.69

测试集中,模型在正常图像上的 MAE 为 1.32, RMSE 为 1.58, PCN 在正常图像上的计数误差较小;在光照不良和透视失真的情况下,MAE 分别为 1.47、1.79, RMSE 分别为 1.95、2.16,说明光照不良与透视失真对模型影响较小,PCN 所具有的多尺度感知结构可以对光照不良与透视失真猪只图像进行较准确的预测;在猪只数量较多和堆叠严重的情况下,模型的 MAE 为 2.35、2.23, RMSE 为 2.82、2.69,说明猪只数量较多和堆叠严重对模型影响较大,但仍然可以计算出较合理的猪只数量范围。模型对复杂场景中的猪只图像计数适应性较强,在真实猪场中的应用是可行的。

3.5.2 不同网络结构对比分析

PCN 在人群计数网络 CSRNet 的基础上添加了 3 层空间金字塔结构,并改进了后端网络的膨胀率。为了验证改进网络结构在猪只计数中的可行性,对不同网络结构进行了测试比对,部分网络结构测试结果如表 3 所示。

对比不同金字塔尺寸的结构,如表 3 中网络 1、2 所示。网络 1 不使用空间金字塔结构,与网络 2 和 PCN 的计数结果相比,使用空间金字塔结构可获

表 3 不同网络结构计数性能比较
Tab.3 Comparison of network counting performance of different structures

网络	金字塔块尺寸	后端网络膨胀率	MAE	RMSE
网络 1		{1,1,2,2,4,4}	2.86	3.67
网络 2	{1,3,6}	{1,1,2,2,4,4}	2.52	3.25
网络 3	{1,2,4}	{2,2,2,2,2,2}	2.23	2.89
网络 4	{1,2,4}	{2,2,4,4,8,8}	2.76	3.73
PCN	{1,2,4}	{1,1,2,2,4,4}	1.74	2.28

得多尺度感知信息,降低模型的计数误差,提高模型鲁棒性。网络 2 使用的金字塔块尺寸为 {1,3,6},与使用块尺寸 {1,2,4} 的 PCN 相比,PCN 优于网络 2,说明提取不同尺寸的空间特征对模型计数的精准性带来影响,中间层需要选取合适的空间金字塔块尺寸。

对比不同的后端网络膨胀率,网络 3 后端网络与 CSRNet 的后端网络相同,6 层网络膨胀率均为 2,与 PCN 相比后端网络感受野更小。网络 4 膨胀率为 {2,2,4,4,8,8},进一步扩大了后端网络感受野。PCN 与网络 3、4 的结果相比,说明在一定程度上扩大后端网络的感受野有助于后端网络进一步提取特征。

通过多次测试与验证,选用金字塔块尺寸为 {1,2,4}、后端网络膨胀率为 {1,1,2,2,4,4} 作为 PCN 的网络结构。在综合考虑空间金字塔块尺寸与后端网络膨胀率的情况下,PCN 的 MAE 为 1.74、RMSE 为 2.28,与表 3 中其他网络相比误差较小、鲁棒性更强。

3.5.3 不同计数模型对比分析

将 PCN 与人群计数网络 MCNN^[11]、CSRNet^[21]和改进 Counting CNN 的猪只计数网络^[7]在测试集上进行性能对比,结果如表 4 所示。

表 4 不同计数方法性能比较
Tab.4 Performance comparison of different counting methods

网络	MAE	RMSE	单幅图像平均识别时间/s
MCNN	10.39	12.83	0.046
CSRNet	3.06	4.42	0.104
改进 Counting CNN	4.27	6.48	0.059
PCN	1.74	2.28	0.108

通过对网络结果进行测试,PCN 的准确率优于 MCNN、CSRNet 和改进 Counting CNN 网络,MAE 和 RMSE 分别为 1.74 和 2.28;虽然单幅图像平均识别时间比其他网络长,但也仅为 0.108 s,满足实时处理要求。

将 PCN 与 MCNN 的计数结果进行比较,使用大、中、小 3 种不同尺度的卷积核,以获得不同的感受野。PCN 前后端网络均使用 3×3 的小尺寸卷积核,通过中间层获取多尺度特征,后端网络使用膨胀卷积来获取更大的感受野。PCN 为更深层的网络结构,参数数量更多,MAE 比 MCNN 减小了 8.65, RMSE 比 MCNN 减小了 10.55。试验结果表明,MCNN 的结构在猪只数据集上效果并不突出,使用深层小卷积核的 PCN 准确率与 MCNN 相比有显著提升。

与 CSRNet 的计数结果相比,PCN 增加了空间金字塔结构作为多尺度感知中间层,并将不同尺度的特征融合,能够感知图像中复杂空间信息。与 CSRNet 相比,PCN 的 MAE 降低了 1.32,计数误差更小,模型计数精度优于 CSRNet, RMSE 降低了 2.14,拥有更强的鲁棒性。虽然 PCN 增加了多尺度感知中间层,单幅图像平均识别时间较长,但较 CSRNet 仅增加了 0.004 s,算法实时性仍然在可接受范围。

与改进 Counting CNN 网络相比,PCN 所采用的前中后端结构均使用小尺寸的卷积核,有更深的网络结构,MAE 比改进 Counting CNN 降低了 2.53, RMSE 降低了 4.20,在计数准确性与鲁棒性上均有所提升,更加适用于高密度猪只计数的情况。

综上所述,PCN 性能明显优于 MCNN、CSRNet 和改进 Counting CNN 网络,在保证图像识别实时性

要求的基础上,在高密度猪只图像上得到了更准确的计数结果。

4 结论

(1)搭建了基于多尺度感知的猪只计数网络,在人群计数网络 CSRNet 的基础上设计了用于猪只计数的网络 PCN。PCN 采用 10 层 VGG16 前端网络提取基础特征,增加了空间金字塔结构作为多尺度感知中间层,从而可对图像中变化的空间尺度特征进行提取与融合,并改进了原网络的后端网络膨胀率,在保持图像分辨率的同时扩大了网络的感受野。PCN 网络可生成高质量的预测密度图,通过对密度图进行积分得到预测的猪只数量,在获取猪只空间分布信息的同时可以有效获取高密度猪群图像中的猪只数量。

(2)与 MCNN、CSRNet 和改进 Counting CNN 的猪只计数网络相比,PCN 的 MAE 和 RMSE 分别为 1.74 和 2.28,MAE 比其他 3 种网络分别低 8.65、1.32 和 2.53, RMSE 分别降低了 10.55、2.14 和 4.20,在高密度猪只图像计数上取得了更好的效果,具有较高的准确性与鲁棒性;模型的单幅图像平均识别时间为 0.108 s,满足算法实时性需求。在数量较多、光照不良、透视失真、堆叠严重等条件下,PCN 仍然可以得到理想的计数结果,在复杂场景中具有较好的适应性。

参 考 文 献

[1] 杨艳,滕光辉,李保明,等.基于计算机视觉技术估算种猪体重的应用研究[J].农业工程学报,2006,22(2):127-131. YANG Yan, TENG Guanghui, LI Baoming, et al. Measurement of pig weight based on computer vision[J]. Transactions of the CSAE, 2006, 22(2):127-131. (in Chinese)

[2] 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.

[3] 张天昊,梁炎森,何志毅.图像识别计数在储备生猪统计的应用[J].计算机应用与软件,2016,33(12):173-178. ZHANG Tianhao, LIANG Yansen, HE Zhiyi. Applying image recognition and counting to reserved live pigs statistics[J]. Computer Applications and Software, 2016, 33(12): 173-178. (in Chinese)

[4] 高云,郭继亮,黎焯,等.基于深度学习的群猪图像实例分割方法[J/OL].农业机械学报,2019,50(4):179-187. GAO Yun, GUO Jiliang, LI Xuan, et al. Instance-level segmentation method for group pig images based on deep learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4):179-187. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190420&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.04.020. (in Chinese)

[5] 杨阿庆,薛月菊,黄华盛,等.基于全卷积网络的哺乳母猪图像分割[J].农业工程学报,2017,33(23):219-225. YANG Aqing, XUE Yueju, HUANG Huasheng, et al. Lactating sow image segmentation based on fully convolutional networks[J]. Transactions of the CSAE, 2017, 33(23): 219-225. (in Chinese)

[6] 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.

[7] TIAN M X, GUO H, CHEN H, et al. Automated pig counting using deep learning[J]. Computers and Electronics in Agriculture, 2019, 163(7): 104840.

[8] 姜海燕,徐灿,陈尧,等.基于田间图像的局部遮挡小尺寸稻穗检测和计数方法[J/OL].农业机械学报,2020,51(9):152-162. JIANG Haiyan, XU Can, CHEN Yao, et al. A detecting and counting method for small-sized and occluded rice panicles based on in-field images[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9):152-162. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200918&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.09.018. (in Chinese)

- [9] OÑORO-RUBIO D, LÓPEZ-SASTRE R J. Towards perspective-free object counting with deep learning [C] // European Conference on Computer Vision, 2016:615 – 629.
- [10] ALAHMARI S S, GOLDFOF D, HALL L, et al. Automated cell counts on tissue sections by deep learning and unbiased stereology[J]. Journal of Chemical Neuroanatomy, 2019,96:94 – 101.
- [11] ZHANG Y Y, ZHOU D, CHEN S Q, et al. Single-image crowd counting via multi-column convolutional neural network[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:589 – 597.
- [12] CHENG Z Q, LI J X, DAI Q, et al. Learning spatial awareness to improve crowd counting [C] // IEEE International Conference on Computer Vision, 2019:6152 – 6161.
- [13] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: an evaluation of the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012,34(4):743 – 761.
- [14] SINDAGI V A, PATEL V M. A survey of recent advances in CNN-based single image crowd counting and density estimation [J]. Pattern Recognition Letters, 2018,107:3 – 16.
- [15] XU B L, QIU G P. Crowd density estimation based on rich features and random projection forest [C] // IEEE Winter Conference on Applications of Computer Vision, 2016:1 – 8.
- [16] IDREES H, SALEEMI I, SEIBERT C, et al. Multi-source multi-scale counting in extremely dense crowd images [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013:2547 – 2554.
- [17] LEMPISKY V, ZISSERMAN A. Learning to count objects in images [C] // International Conference on Neural Information Processing Systems, 2010:1324 – 1332.
- [18] WANG Y, ZOU Y X. Fast visual object counting via example-based density estimation [C] // IEEE International Conference on Image Processing, 2016:3653 – 3657.
- [19] ZHANG Y M, ZHOU C L, CHANG F L, et al. Multi-resolution attention convolutional neural network for crowd counting[J]. Neurocomputing, 2019, 329: 144 – 152.
- [20] BOOMINATHAN L, KRUTHIVENTI S S S, VENKATESH B R. Crowdnet: a deep convolutional network for dense crowd counting [C] // Proceedings of the 2016 ACM on Multimedia Conference, 2016: 640 – 644.
- [21] LI Y H, ZHANG X F, CHEN D M. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes [C] // Conference on Computer Vision and Pattern Recognition, 2018.
- [22] KANG D, MA Z, CHAN A B. Beyond counting: comparisons of density maps for crowd analysis tasks-counting, detection, and tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019,29(5):1408 – 1422.
- [23] CHEN Y, GAO C, SU Z, et al. Scale-aware rolling fusion network for crowd counting [C] // IEEE International Conference on Multimedia and Expo, 2020: 1 – 6.
- [24] PHAM V Q, KOZAKAYA T, YAMAGUCHI O, et al. Count forest: co-voting uncertain number of targets using random forest for crowd density estimation [C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 3253 – 3261.
- [25] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:3431 – 3440.
- [26] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C] // International Conference on Learning Representations, 2015.
- [27] SHI M J, YANG Z H, XU C, et al. Revisiting perspective information for efficient crowd counting [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019:7279 – 7288.
- [28] SAM D B, SURYA S, VENKATESH B R. Switching convolutional neural network for crowd counting [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4031 – 4039.
- [29] SINDAGI V A, PATEL V M. Generating high-quality crowd density maps using contextual pyramid CNNs [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:1861 – 1870.
- [30] LIU W Z, SALZMANN M, FUA P. Context-aware crowd counting [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019:5099 – 5108.
- [31] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [C] // European Conference on Computer Vision, 2014.
- [32] 高云, 廖慧敏, 黎焯, 等. 基于双金字塔网络的 RGB-D 群猪图像分割方法 [J/OL]. 农业机械学报, 2020, 51(7):36 – 43. GAO Yun, LIAO Huimin, LI Xuan, et al. RGB-D segmentation method for group piglets images based on the double-pyramid network [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7):36 – 43. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200705&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2020. 07. 005. (in Chinese)
- [33] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [C] // International Conference on Learning Representations, 2016.
- [34] 彭贤, 彭玉旭, 汤强, 等. 基于单列多尺度卷积神经网络的人群计数 [J]. 计算机科学, 2020, 47(4):150 – 156. PENG Xian, PENG Yuxu, TANG Qiang, et al. Crowd counting based on single-column multi-scale convolutional neural network [J]. Computer Science, 2020, 47(4):150 – 156. (in Chinese)