

基于注意力机制和可变形卷积的鸡只图像实例分割提取

方 鹏¹ 郝宏运¹ 李腾飞² 王红英¹

(1. 中国农业大学工学院, 北京 100083; 2. 北京城市学院北京 3D 打印研究院, 北京 100083)

摘要: 为提高鸡只个体轮廓分割提取的精度和准确度,实现基于机器视觉技术的鸡只行为、健康、福利状态监测等精准畜牧业管理,保证相关监测技术及决策的可靠性,针对叠层笼养环境下肉鸡图像的实例分割和轮廓提取问题,提出一种优化的基于 Mask R-CNN 框架的实例分割方法,构建了一种鸡只图像分割和轮廓提取网络,对鸡群图像进行分割,从而实现鸡只个体轮廓的提取。该网络以注意力机制、可变形卷积的 41 层深度残差网络 (ResNet) 和特征金字塔网络 (Feature pyramid networks, FPN) 相融合为主干网络,提取图像特征,并经区域生成网络 (Region proposal networks, RPN) 提取感兴趣区域 (ROI),最后通过头部网络完成鸡只目标的分类、分割和边框回归。鸡只图像分割试验表明,与 Mask R-CNN 网络相比,优化后网络模型精确率和精度均值分别从 78.23%、84.48% 提高到 88.60%、90.37%,模型召回率为 77.48%,可以实现鸡只轮廓的像素级分割。本研究可为鸡只福利状态和鸡只健康状况的实时监测提供技术支撑。

关键词: 肉鸡; 实例分割; 轮廓提取; 可变形卷积神经网络; 注意力机制

中图分类号: TP391.41; S818.9 文献标识码: A 文章编号: 1000-1298(2021)04-0257-09

OSID:



Instance Segmentation of Broiler Image Based on Attention Mechanism and Deformable Convolution

FANG Peng¹ HAO Hongyun¹ LI Tengfei² WANG Hongying¹

(1. College of Engineering, China Agricultural University, Beijing 100083, China

2. Beijing 3D Printing Research Institute, Beijing City University, Beijing 100083, China)

Abstract: Segmentation and extraction of birds contour is the premise of precision livestock farming management, such as behavior, health, welfare status monitoring based on machine vision technology. The precision and accuracy of image segmentation directly affect the reliability of relevant monitoring technology and decision-making. An instance segmentation approach based on Mask R-CNN deep learning framework was proposed to solve broiler instance segmentation and contour extraction problems in stacked-cage henhouse. Furthermore, a broiler image segmentation and contour extraction network was constructed to segment broiler images and realize birds individual contour extraction. In this network, totally 41 layers deep residual network (ResNet) based on attention mechanism and deformable convolution was integrated with feature pyramid networks (FPN) as the backbone network to extract the image features, and regions of interest were extracted by region proposal networks. Finally, target classification, segmentation and box regression were realized through network heads. Broiler image segmentation experiment showed that compared with Mask R-CNN network, the average precision and mean accuracy of the optimized network were improved from 78.23% and 84.48% to 88.60% and 90.37%, respectively, and the recall rate of the model was 77.48%, which can realize the pixel level segmentation of chicken contour. The research result can provide technical support for the real-time monitoring of birds welfare and health status.

Key words: broiler; instance segmentation; contour extraction; deformable convolution neural network; attention mechanism

收稿日期: 2020-11-04 修回日期: 2021-01-18

基金项目: 国家重点研发计划项目 (2017YFE0122200)

作者简介: 方鹏 (1993—), 男, 博士生, 主要从事畜禽养殖技术与装备研究, E-mail: fangpeng@cau.edu.cn

通信作者: 王红英 (1966—), 女, 教授, 博士生导师, 主要从事畜禽养殖技术与装备研究, E-mail: hongyingw@cau.edu.cn

0 引言

人类对畜牧业产品的消费需求逐年上升,同时,畜牧业所需的生产资料(土地、水资源、劳动力等)却不断萎缩^[1],生产力与社会需求的矛盾日益凸显。在这种情况下,精准畜牧业^[2-3]理念以其可持续、高效、低耗的特点应运而生。获取动物的个体行为、健康、福利信息对精准畜牧业管理决策十分重要^[4-5]。机器视觉技术广泛应用于动物监测中,而基于视觉的监测技术的前提是实现对动物图像的提取和分割,只有对动物轮廓信息进行精准分割,才能对动物个体进行生长评估^[6]、体况评价和行为分析^[7-8](如发情行为、产前行为)等方面的研究。

机器视觉技术具有采集速度快、识别精度高的特点,同时更具有无接触、对动物应激小的优势,在鸡只体况监测^[9-12]、鸡只行为识别^[13-14]和鸡只福利状态监测^[15-16]等方面都有广泛的应用。相关研究在很大程度上依赖于图像的分割,显然,图像分割的准确性和精度对基于机器视觉的鸡只监测技术至关重要。然而,传统的基于颜色空间^[17]、帧差或者光流^[18]的分割方法难以在复杂养殖环境中实现高精度、高鲁棒性的图像分割。

近年来,具有很强特征学习能力的深度卷积神经网络在计算机视觉领域得到了广泛应用。通过大量经过人工标注的图像数据的训练,神经网络可以充分提取携带丰富空间和语义信息的图像特征,并将其用于图像分割,其分割效果良好^[19]。作为一种像素级识别分割目标轮廓的目标检测算法,实例分割最早在2014年提出,并迅速得到发展,先后出现了DeepMask^[19]、SharpMask^[20]等实例分割算法,但均存在精度不高、模型泛化性不够的缺点。文献[21]提出的Mask R-CNN框架在模型泛化能力、分类精度和分割精度上均有优异的表现。研究表明,通过深度卷积神经网络实现养殖环境下鸡只轮廓的提取和分割是可行的。

为了实现叠层笼养环境下鸡只图像的高精度分割和轮廓提取,本文提出一种基于Mask R-CNN的鸡只图像分割和轮廓提取方法,先对原始图像进行增强,以提高图像品质,随后进行鸡只图像的分割和轮廓的提取。

1 图像采集与图像增强处理

以叠层笼养环境下的白羽肉鸡为研究对象,进行识别分割试验。于2019年7月9—11日在山东省烟台市蓬莱区民和牧业股份有限公司肉鸡养殖场采集了27~29日龄的肉鸡图像。试验鸡舍为一栋

全封闭的8层叠层笼养肉鸡舍,舍内平均光照强度为8 lx,共饲养罗斯308肉鸡114 240只。

图像采集所用设备主要包括:Sony XCG-240C型彩色数字相机,分辨率为1 920像素×1 200像素,匹配焦距6 mm的Ricoh FL-CC0614A-2M型定焦镜头;立式三脚架,最大拍摄高度可达2.2 m。拍摄时相机参数设定为:采集帧率3 f/s,曝光时间80 ms。

将工业相机安装在立式三角架上,固定相机的工作距离、拍摄高度和拍摄角度,在笼门前方位置对笼内鸡群进行拍摄。将采集的视频按帧提取成静态图像,人工挑选出800幅肉鸡图像建立样本集。由于舍内照度较低,采集的图像亮度不够,为便于后续的图像标注和图像特征提取,先对采集的图像进行Retinex增强。同时,为减少计算量,降低模型训练时间,将原图像调整为448像素×256像素。随后,在Labelme图像标注工具中用多边形标注出鸡只的外轮廓,共标注目标11 034个,制作成COCO(Common objects in context)格式的数据集。在样本集中随机选取70%(560幅)图像作为训练集,用于模型的训练,选取20%(160幅)图像作为验证集,用于训练参数的调优,选取10%(80幅)图像作为测试集,用于评估最终模型的识别能力。采集的原图及Retinex增强后的图像如图1所示。

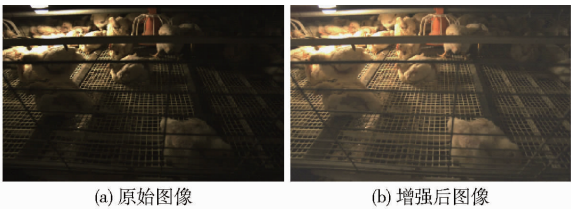


图1 鸡只图像
Fig.1 Broiler images

2 基于Mask R-CNN的鸡只轮廓实例分割模型优化

Mask R-CNN作为一种实例分割算法,集成了目标检测和语义分割两大功能,可以同时完成目标分割、分类和检测任务,且在这3种任务上均有较优的表现,是一种综合性能很优异的实例分割方法。因此,本文选择在Mask R-CNN模型框架下进行鸡只图像实例分割模型的研究,其网络结构如图2所示。

整体上,Mask R-CNN网络分为主干网络、区域生成网络(Region proposal networks, RPN)和头部网络。主干网络主要进行图像特征的提取,区域生成网络主要通过提取的图像特征生成感兴趣区域(ROI),而头部网络则在此基础上完成之后的目标分类、边框回归及掩膜预测生成工作。

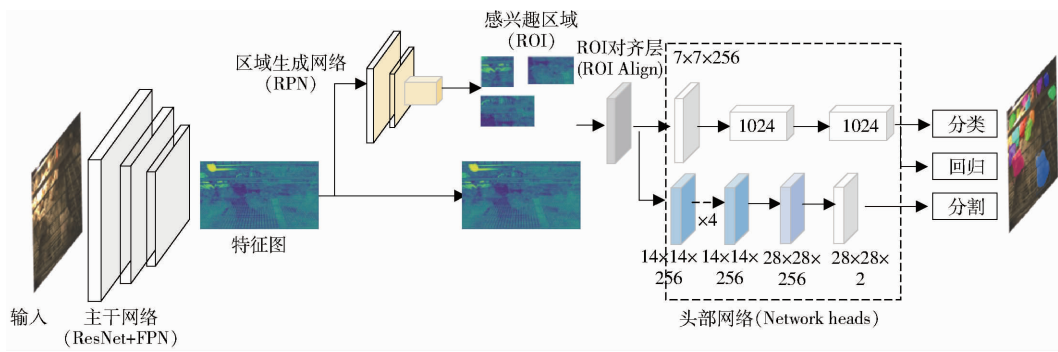


图 2 Mask R-CNN 网络结构图

Fig. 2 Network structure of Mask R-CNN

针对本文待分割图像的特点,在现有 Mask R-CNN 网络基础上进行了调整和优化,构建了一种鸡只图像分割和轮廓提取网络。主要对主干网络做出 3 点优化:①调整卷积层结构。②构建基于注意力机制的卷积层。③引入可变形卷积。

2.1 主干网络卷积层结构调整优化

目前,通用的 Mask R-CNN 网络以深度残差网络^[22] (ResNet) 和特征金字塔网络 (Feature pyramid networks, FPN) 相结合的方式作为主干网络,负责输入图像的特征提取。ResNet 网络作为一种深层网络,可以有效解决常规网络堆叠到一定深度时出现的梯度弥散问题,通过深层次的网络达到较好的特征提取效果,其基本结构为残差模块 (Residual block),通过模块的堆叠,使网络达到较大深度。

现有研究表明,浅层网络的感受野较小,能够捕获更多的图像细节,提升检测的精度;而深层网络输出的特征更加抽象,更加关注图像的语义信息,有利于目标的检出^[23]。本文研究对象为笼养状态下的鸡只,目标单一,类别少,目标检出相对容易。因此对现有主干网络卷积层数量和残差学习模块堆叠方式进行调整,适当削减卷积层的数量,增加浅层网络

深度,减小深层网络深度,使网络更加关注图像的细节信息。网络由原来的 101 层卷积层减少为 41 层,降低网络计算量;同时网络第 3 阶段卷积层数量从 12 层增加为 15 层,第 4 阶段卷积层数量从 69 层减少为 9 层,第 5 阶段卷积层数量从 9 层减少为 6 层,降低图像特征损耗。调整后的 ResNet 网络结构如图 3 所示。图中只包含卷积层和池化层,每个卷积层后还有批量正则化层和激活层未画出;第 1~5 阶段分别为 ResNet 网络的 5 个阶段;2x、4x 为模块重复次数;Conv(64, 256, $k=(1,1)$, $s=1$, $p=0$) 中, k 为卷积核尺寸, s 为滑动步长, p 为填充像素数, 256 为卷积核通道数, 64 为上一层卷积层输出的通道数,其余卷积层类似。主干网络由残差学习模块堆叠而成,残差学习模块首端和末端为 1×1 卷积核,中间为 3×3 卷积核。在网络每个阶段的第 1 个残差模块,除了 3 个卷积层的串联,输入和输出之间还通过一个卷积层旁路相连,以增加输入特征图的通道数,便于和输出特征图融合,而后面接的残差学习模块输入和输出特征图的通道数一致,故可以不通过卷积层升高维度而直接进行加操作。这种结构可以有效降低特征损耗,提升模型训练效果。

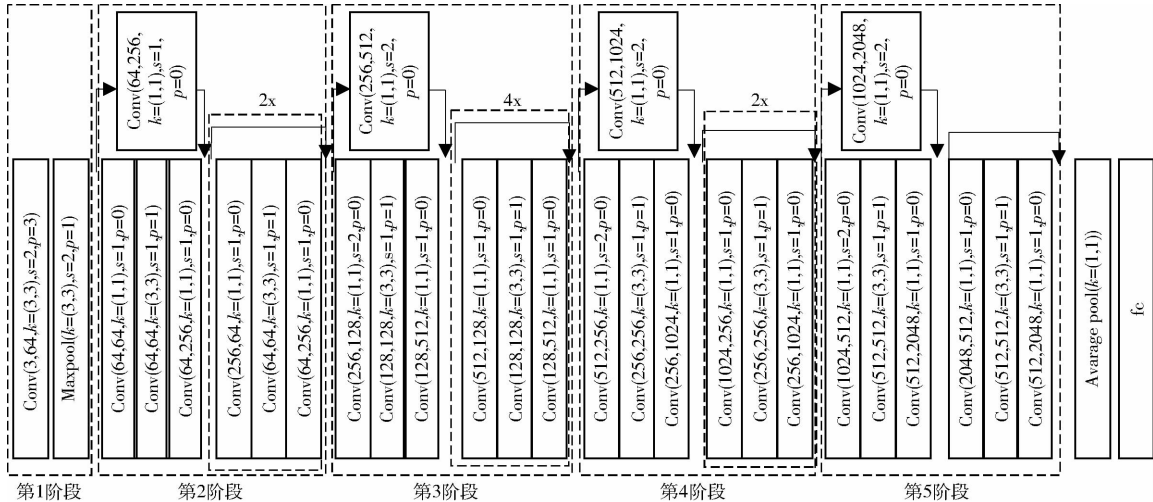


图 3 优化后的 ResNet 网络结构

Fig. 3 Optimized ResNet network structure

FPN^[23]网络结构如图4所示。输入图像自下而上分别经过ResNet网络5个阶段的处理,输出5种不同尺度的特征图(C1~C5)。输出的特征图分别通过一个1×1的卷积操作后与自上而下的上采样操作生成的特征图进行融合,生成特征图M5、M4、M3和M2,随后经过3×3的卷积操作消除上采样的混叠效应,生成特征图P5、P4、P3和P2,P5经过一次下采样操作生成P6,特征图P2、P3、P4、P5和P6分别独立地输入到RPN网络,生成若干个感兴趣区域(ROI)。FPN网络将ResNet网络每一阶段输出的特征图融合,既利用了高层特征图的强语义信息,又利用了低层特征图的强空间位置信息,大大提升了主干网络的特征提取能力。

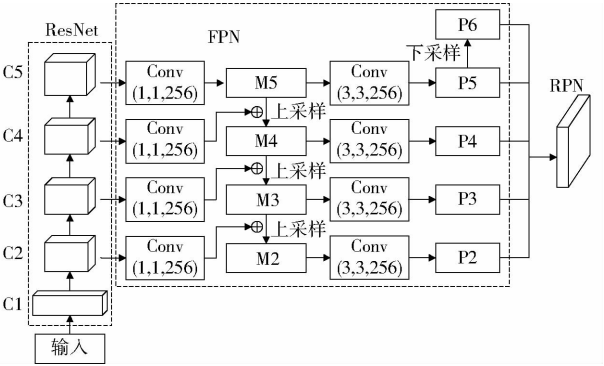


图4 FPN网络结构图
Fig. 4 FPN network structure diagram

2.2 基于注意力机制的卷积层构建

注意力机制最早由MNIH等^[24]提出并引入图像分类领域,随后在机器学习领域迅速发展,众多应用于不同领域的注意力模型相继提出。注意力机制模仿人类视觉系统的处理机制,人类在进行视觉信息处理时会自动过滤不重要的信息,而把更多注意力资源用于需要重点关注的目标区域,大大提高了视觉信息处理效率与准确性。而当注意力机制引入基于神经网络的图像处理领域时,其本质是一种资源的重新分配机制,即对不同重要程度的信息赋予不同的权重,大大提高神经网络效率,用很小的计算量换取网络性能的显著提升。

为提升鸡只轮廓实例分割模型网络性能,本文在ResNet网络中引入注意力机制。在原有网络结构中添加1个通道注意力模块和1个空间注意力模块^[25]。

通道注意力模块结构如图5a所示,输入的特征图经过1个全局最大池化层和平均池化层后,分别得到1个通道描述,随后这2个通道描述送入1个2层的神经网络,得到2个特征向量,2个特征向量通过逐元素累加的方式合并成1个,通过1个激活函数输出通道权重系数,权重系数与输入特征相乘即得到新的特征图。空间注意力模块结构如图5b所示。

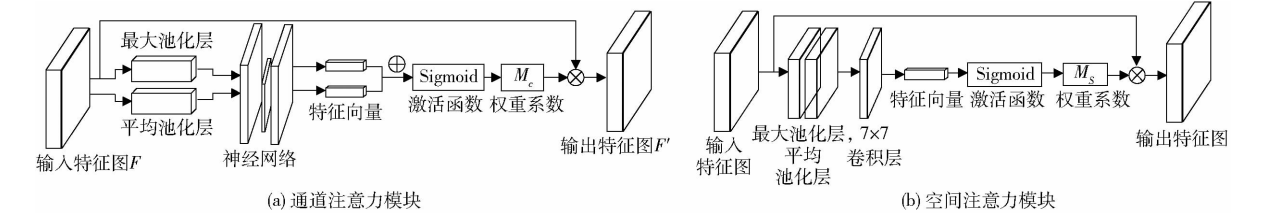


图5 注意力模块
Fig. 5 Attention module

本文在ResNet网络的每一个残差模块中串联插入1个通道注意力模块和1个空间注意力模块。卷积运算通过将跨通道和空间信息融合在一起提取信息特征,既考虑了不同通道像素的重要性,又考虑了同一通道不同位置像素的重要性。其在ResNet网络残差学习模块的位置如图6所示。每个残差学习模块中,在第2个1×1卷积层之后串联接入1个通道注意力模块和1个空间注意力模块。

2.3 引入可变形卷积层

在图像检测任务(目标检测、实例分割)中,通常需要网络对同一目标的不同姿态,如位置的偏移、角度的旋转及尺度的变化都具有相同的识别能力。然而,受限于卷积操作本身的固定性,卷积神经网络并不具有尺度不变性和旋转不变性,基本没有对目

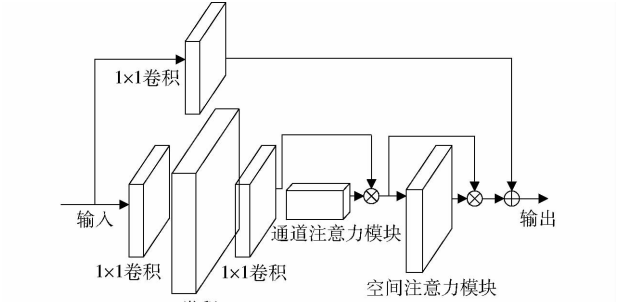


图6 残差模块中添加的注意力模块示意图
Fig. 6 Attention module integrated with residual module

标几何形变的适应能力^[26]。实际中,神经网络对目标变化的适应能力几乎完全来自于数据本身的多样性。现有研究表明,标准卷积中的规则格点采样是导致网络难以适应几何形变的主要原因,为此DAI

等^[27]提出了一种可变形卷积网络,用可变形卷积取代神经网络中的标准卷积,众多研究表明,该方法在目标检测领域表现优异^[28-30]。

标准的卷积操作具有非常固定的几何结构,很难与目标复杂的外形相匹配。而可变形卷积拥有可变化、不规则的形状,感受野灵活多变,可以很好适应不同尺寸、外形的目标。在标准的二维卷积过程中,对于输出特征图中的每一个位置 P_0 ,其特征值 $y(P_0)$ 计算式为

$$y(P_0) = \sum_{P_n \in R} w(P_n) x(P_0 + P_n) \tag{1}$$

式中 $w(P_n)$ ——该采样位置的卷积核权重
 $x(P_0 + P_n)$ ——采样位置的输入特征值
 P_n ——感受野区域中的所有采样位置
 R ——感受野区域

而在可变形卷积中

$$y(P_0) = \sum_{P_n \in R} w(P_n) x(P_0 + P_n + \Delta P_n) \tag{2}$$

式中 $x(P_0 + P_n + \Delta P_n)$ ——采样偏移位置的输入特征值
 ΔP_n ——采样点位置的偏移量

可见,可变形卷积就是在传统的卷积操作上加入了一个采样点位置的偏移量,使卷积变形为不规则卷积,从而拥有更大、更灵活的感受野。

为分割出图像中鸡只的轮廓像素,本文在 ResNet 网络中加入可变形卷积层,通过引入偏移量,增大感受野,同时使感受野可以适应不同尺寸、形状的鸡只轮廓,达到更好的分割效果。将 ResNet 网络的第 3、4、5 阶段的 3×3 卷积调整为可变形卷积,而第 2 阶段保留为标准卷积层,以减小引入可变形卷积带来的网络参数数量的增加对网络负荷的影响。

3 鸡只轮廓识别分割试验

3.1 试验条件及配置

试验在 Ubuntu 18.04 系统下进行,试验所用机器配置:处理器为 Intel(R) Core(TM) i7-9700K,主频 3.6 GHz,内存 16 GB,显卡为 NVIDIA GeForce RTX 2080(16 GB),使用 GPU 加速计算,采用 Python 作为编程语言,选择 Pytorch 框架来实现网络模型的搭建、训练和调试。训练集为 560 幅图像,验证集为 160 幅图像,测试集为 80 幅图像。

使用 Torchvision 视觉库中的 ResNet 预训练模型作为初始输入权重,采用随机梯度下降法对鸡只轮廓分割网络进行训练。设置学习率为 0.001,采用热身策略,初始学习率为 0.000 18,采用线性增加策略,训练 5 个 epochs 后增加到 0.001;动量为 0.9,

权值衰减系数为 0.000 1,训练迭代 100 个 epochs,每 5 个 epochs 保存一个权重,取精度最高的模型为最终模型。

3.2 试验结果与分析

3.2.1 不同模型检测结果分析

采用不同的主干网络进行鸡只轮廓分割试验:
①现有的 ResNet 101 网络。
②调整现有网络卷积层结构后得到的 ResNet 41 网络。
③ResNet 41 网络中添加注意力机制(简称为 ResNet 41 + cbam)。
④ResNet 41 网络中添加注意力机制并将部分卷积层替换成可变形卷积(简称为 ResNet 41 + cbam + dcn)。
⑤ResNet 50 网络中添加注意力机制和可变形卷积层(简称为 ResNet 50 + cbam + dcn)。

模型对鸡只图像的分割结果如图 7 所示。由图 7 可知,分割模型可较准确地将鸡只轮廓之间的粘连区域分割开,可实现笼养鸡只轮廓的提取分割。

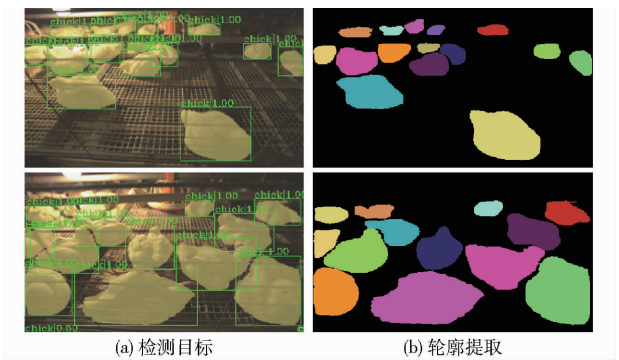


图 7 鸡只轮廓识别分割模型分割结果
Fig.7 Broiler image recognition and segmentation results of model

以准确率 A (Accuracy)、召回率 R (Recall) 和精确率 P (Precision)、平均检测时间 t 作为评价指标,衡量分割模型的性能。

不同模型的性能如表 1 所示。由表 1 可知,将主干网络卷积层数从 101 层减小为 41 层,模型各指标未有显著降低,说明原有的 101 层卷积网络对本文研究的鸡只轮廓分割提取任务有较大冗余,降低其卷积层数不影响模型性能,但可以大幅降低运算量,其单幅图像检测时间从 0.32 s 减小为 0.18 s,降低了 44%。在 ResNet 41 网络的基础上引入注意力机制,模型各指标性能有较大提升,其精确率和准确率分别从 77.01%、82.34% 大幅提升至 85.49%、88.35%,召回率略有下降,而检测时间却没有显著增加,注意力机制的引入使模型用很小的资源开销换取了较大的性能提升。在引入注意力机制的基础上,将部分 3×3 卷积层调整为可变形卷积层,模型性能较优化前有一定程度的提升,精确率和准确率分别从 85.49%、88.35% 提升到 88.60%、90.37%,

与现有的 Mask R - CNN 网络相比,其精确率和准确率分别提高了 10.37、5.89 个百分点。但同时,可变形卷积的引入增加了模型的参数量,其检测时间从 0.24 s 增至 0.41 s,对模型的实时性有所影响。在 ResNet 50 网络中引入注意力机制并添加可变形卷积层,模型在验证集上的精确率和准确率分别为 87.23%、89.80%,均略低于在 ResNet 41 网络中引入注意力机制及添加可变形卷积层的模型性能,说明本文设计的 41 层网络结构有利于提升分割网络的性能。

表 1 不同网络性能对比

Tab.1 Performance comparison of different networks

主干网络	精确率/%	准确率/%	召回率/%	平均检测时间/s
ResNet 101	78.23	84.48	80.12	0.32
ResNet 41	77.01	82.34	82.53	0.18
ResNet 41 + cbam	85.49	88.35	78.12	0.24
ResNet 41 + cbam + dcn	88.60	90.37	77.48	0.41
ResNet 50 + cbam + dcn	87.23	89.80	79.59	0.47

3.2.2 不同模型损失曲线分析

对上述 5 个网络训练过程中的损失函数变化进行分析。损失函数衡量的是模型训练过程中预测值和真实值之间的差异变化。在基于 Mask R - CNN 网络的鸡只轮廓实例分割模型的训练中,损失函数 L 主要由分类损失、检测损失和分割损失 3 部分组成,定义式^[21,31]为

$$L = L_{cls} + L_{bbox} + L_{mask}$$
 (3)

式中 L_{cls} ——模型 的分类损失
 L_{bbox} ——检测损失(边框回归损失)
 L_{mask} ——分割损失
分类损失 L_{cls} 计算式为

$$L_{cls} = -\frac{1}{N_{cls}} \sum_i \text{lb}(p_i^* p_i + (1 - p_i^*) (1 - p_i))$$
 (4)

式中 N_{cls} ——类别数量
 p_i ——目标被预测为正样本的概率
 p_i^* ——模型判定值,当区域判定为正样本时, $p_i^* = 1$,否则为 0
检测损失 L_{bbox} 计算式为

$$L_{bbox} = \frac{1}{N_{reg}} \sum_i p_i^* R(t_i - t_i^*)$$
 (5)

其中

$$R(t_i - t_i^*) = \begin{cases} 0.5 (t_i - t_i^*)^2 & (|t_i - t_i^*| < 1) \\ |t_i - t_i^*| - 0.5 & (|t_i - t_i^*| \geq 1) \end{cases}$$
 (6)

式中 $R(t_i - t_i^*)$ ——平滑损失函数

N_{reg} ——特征图的像素值
 t_i ——预测边框的坐标向量
 t_i^* ——真实边框的坐标向量

分割损失 L_{mask} 计算式为

$$L_{mask} = \text{Sigmoid}(C_k)$$
 (7)

式中 C_k ——第 k 类目标

网络对于每一个 ROI 都有 $k \times m^2$ 维度的输出, k 为类别数,共输出 k 个分辨率为 $m \times m$ 的二值掩膜。对于第 k 类目标(C_k), L_{mask} 定义为对掩膜中的每一个像素执行 Sigmoid 函数得到的平均二值交叉熵损失。

图 8 为以 ResNet 41 为主干网络的鸡只轮廓分割模型训练过程中检测损失、分类损失、分割损失和总损失的变化。各损失均在训练开始的很短时间内下降到较低值,随后随着迭代步数的增加缓慢下降,训练过程中各损失曲线波动较小,模型收敛较好,说明各超参数配置较为合理。当迭代步数达到 10 000 次(100 个 epochs)左右时,各损失均趋于稳定,不再持续下降。

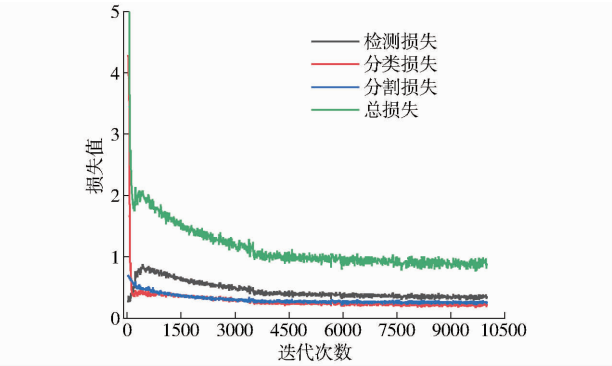


图 8 ResNet 41 网络训练损失曲线

Fig. 8 Training loss curves of ResNet 41 network

图 9 为不同网络的总损失随迭代步数的变化情况。5 个网络的总损失均随着网络的迭代逐步下降,最终趋于稳定,模型收敛。ResNet 101 和 ResNet 41 网络的损失曲线在训练后期基本重合,最终训练损失分别为 0.85 和 0.87 左右,一定程度上说明 2 个模型具有相似的性能表现。而 ResNet 41 + cbam 网络和 ResNet 41 + cbam + dcn 网络最终训练损失分别稳定在 0.63 和 0.31 左右,较优化之前的网络有较明显的下降,将部分卷积层替换成可变形卷积后,模型性能有一定提升。ResNet 50 + cbam + dcn 的网络最终训练损失在 0.43 左右,略高于 ResNet 41 + cbam + dcn 网络,ResNet 41 + cbam + dcn 网络在所有试验网络中性能最优。

3.2.3 不同模型输出特征可视化分析

为更好地解释本文采用的优化方式对模型性能的提升,利用 GRAD - CAM^[32] 对不同网络进行可视

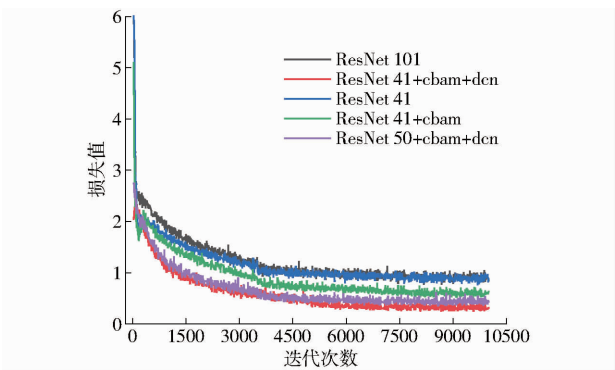


图 9 不同网络训练损失曲线

Fig. 9 Training loss curves of different networks

化分析,通过梯度来衡量卷积层中空间位置的重要性,分别输出不同网络、不同阶段输出特征的类激活热力图(Class activation heat map),可以清楚地显示网络在进行预测时重点关注的图像区域。本文将 5 种网络第 2 阶段和第 5 阶段的特征可视化结果进行了比较,结果如图 10 所示。图中红色区域的范围越大,说明网络提取的特征更多地覆盖到了需要识别

的目标上。由图可知,ResNet 101 网络和 ResNet 41 网络均只有少部分特征覆盖到了鸡只轮廓上,而 ResNet 41 + cbam 网络、ResNet 41 + cbam + dcn 及 ResNet 50 + cbam + dcn 网络提取的特征更好地覆盖了目标对象区域,说明在网络中引入注意力机制确实加强了网络对重点信息的关注程度,提升了网络性能。而可变形卷积使网络具有更大、更灵活的感受野,提升了网络对不同尺寸、不同外形目标的适应能力,提升了网络对鸡只轮廓的分割能力。对比 ResNet 41 + cbam + dcn 网络和 ResNet 50 + cbam + dcn 网络的可视化结果,可以发现 ResNet 41 + cbam + dcn 网络特征提取效果略好。从第 2 阶段和第 5 阶段的特征图可以看出,随着网络的加深,网络提取的特征越来越多地覆盖到了鸡只轮廓上,同时网络也过滤掉了一些不属于鸡只轮廓的特征,提升了模型的检测精度。模型输出特征的可视化分析结果与上文中模型性能指标和训练损失分析结果一致,进一步说明本文对网络进行的优化是有效的。

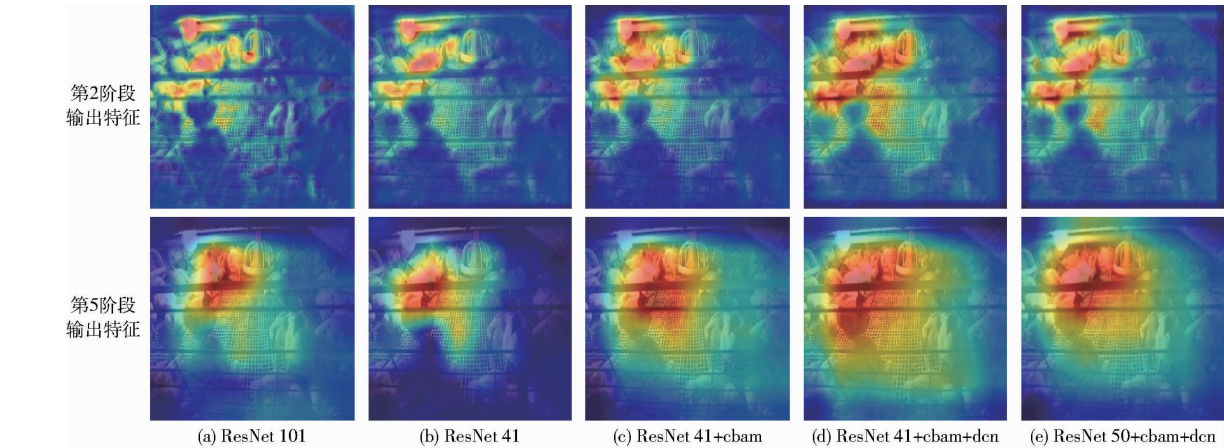


图 10 不同网络不同输出层的热力图

Fig. 10 Heat maps of different networks in different layers

4 结论

(1)以叠层笼养下的肉鸡为研究对象,将 Mask R - CNN 网络引入鸡只轮廓的分割提取中,构建了一种鸡只图像分割和轮廓提取网络。该网络以基于注意力机制、可变形卷积的 41 层深度残差网络和特征金字塔网络相融合的方式作为主干网络,可以实现笼养状态下肉鸡图像的分割和轮廓提取。

(2)优化后的模型在验证集的精确率、准确率

和召回率分别为 88. 60%、90. 37% 和 77. 48%,与现有的 Mask R - CNN 网络相比,其精确率和准确率分别提高了 10. 37、5. 89 个百分点,而单幅图像的检测时间仅增加了 0. 09 s,说明注意力机制和可变形卷积的引入有效提高了网络的综合性能。

(3)特征图可视化分析表明,网络中引入注意力机制和可变形卷积后,网络提取的特征更多地覆盖到鸡只轮廓上,提高了检测精度。本文算法模型对笼养状态下的鸡群分割效果较好,能准确分割提取鸡只个体轮廓。

参 考 文 献

[1] THORNTON P K. Livestock production: recent trends, future prospects[J]. Philosophical Transactions of the Royal Society B—Biological Sciences, 2010, 365(1554): 2853 - 2867.
[2] BERCKMANS D. Precision livestock farming (PLF)[J]. Computers and Electronics in Agriculture, 2008, 62(1): 1.
[3] TULLO E, FONTANA I, GUARINO M. Precision livestock farming: an overview of image and sound labelling[C] // The

- European Conference on Precision Livestock Farming, 2013: 30–38.
- [4] BOS J M, BOVENKERK B, FEINDT P H, et al. The quantified animal: precision livestock farming and the ethical implications of objectification[J]. *Food Ethics*, 2018, 2(1): 77–92.
- [5] 何东健, 刘冬, 赵凯旋. 精准畜牧业中动物信息智能感知与行为检测研究进展[J/OL]. *农业机械学报*, 2016, 47(5): 231–244.
HE Dongjian, LIU Dong, ZHAO Kaixuan. Review of perceiving animal information and behavior in precision livestock farming [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(5): 231–244. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160532&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.05.032. (in Chinese)
- [6] KASHIHA M, BAHR C, OTT S, et al. Automatic weight estimation of individual pigs using image analysis[J]. *Computers and Electronics in Agriculture*, 2014, 107: 38–44.
- [7] 庄晏榕, 余昊桦, 滕光辉, 等. 基于卷积神经网络的大白母猪发情行为识别方法研究[J/OL]. *农业机械学报*, 2020, 51(增刊1): 364–370.
ZHUANG Yanrong, YU Jionghua, TENG Guanghui, et al. Recognition method of large white sow oestrus behavior based on convolutional neural network[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(Supp. 1): 364–370. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2020s143&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.S1.043. (in Chinese)
- [8] 刘忠超, 何东健. 基于卷积神经网络的奶牛发情行为识别方法[J/OL]. *农业机械学报*, 2019, 50(7): 186–193.
LIU Zhongchao, HE Dongjian. Recognition method of cow estrus behavior based on convolutional neural network [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(7): 186–193. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190719&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.019. (in Chinese)
- [9] 毕敏娜, 张铁民, 庄晓霖, 等. 基于鸡头特征的病鸡识别方法研究[J/OL]. *农业机械学报*, 2018, 49(1): 51–57.
BI Minna, ZHANG Tiemin, ZHUANG Xiaolin, et al. Recognition method of sick yellow feather chicken based on head features [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(1): 51–57. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180106&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.01.006. (in Chinese)
- [10] ZHUANG X L, ZHANG T M. Detection of sick broilers by digital image processing and deep learning[J]. *Biosystems Engineering*, 2019, 179: 106–116.
- [11] AYDIN A. Development of an early detection system for lameness of broilers using computer vision[J]. *Computers & Electronics in Agriculture*, 2017, 136:140–146.
- [12] 沈明霞, 李嘉位, 陆明洲, 等. 基于动态多特征变量的黄羽肉鸡跛行状态定量评价方法[J/OL]. *农业机械学报*, 2018, 49(9): 35–44.
SHEN Mingxia, LI Jiawei, LU Mingzhou, et al. Evaluation method of limping status of broilers based on dynamic multi-feature variables[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(9): 35–44. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180904&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.09.004. (in Chinese)
- [13] 劳凤丹, 杜晓冬, 滕光辉. 基于深度图像的蛋鸡行为识别方法[J/OL]. *农业机械学报*, 2017, 48(1): 155–162.
LAO Fengdan, DU Xiaodong, TENG Guanghui. Automatic recognition method of laying hen behaviors based on depth image processing[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(1): 155–162. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170120&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.01.020. (in Chinese)
- [14] 劳凤丹, 滕光辉, 李军, 等. 机器视觉识别单只蛋鸡行为的方法[J]. *农业工程学报*, 2012, 28(24): 157–163.
LAO Fengdan, TENG Guanghui, LI Jun, et al. Behavior recognition method for individual laying hen based on computer vision [J]. *Transactions of the CSAE*, 2012, 28(24): 157–163. (in Chinese)
- [15] PEÑA F A, NORTON T, TULLO E, et al. Real-time monitoring of broiler flock's welfare status using camera-based technology [J]. *Biosystems Engineering*, 2018, 173:103–114.
- [16] KASHIHA M, PLUK A, BAHR C, et al. Development of an early warning system for a broiler house using computer vision [J]. *Biosystems Engineering*, 2013, 116(1): 36–45.
- [17] 毕敏娜, 张铁民, 庄晓霖, 等. 基于色差信息多色彩模型的黄羽鸡快速分割方法[J/OL]. *农业机械学报*, 2016, 47(12):293–298.
BI Minna, ZHANG Tiemin, ZHUANG Xiaolin, et al. Fast segmentation method of yellow feather chicken based on difference of color information in different color models[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2016, 47(12):293–298. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20161236&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.12.036. (in Chinese)
- [18] DAWKINS M S, CAIN R, MERELIE K, et al. In search of the behavioral correlates of optical flow patterns in the automated assessment of broiler chicken welfare[J]. *Applied Animal Behavior Science*, 2013, 145(1–2): 44–50.
- [19] PINHEIRO P O, COLLOBERT R, DOLLAR P. Learning to segment object candidates[C]// *The Conference on Neural Information Processing Systems*, 2015: 1990–1998.
- [20] PINHEIRO P O, LIN T Y, COLLOBERT R, et al. Learning to refine object segments[C]// *European Conference on Computer Vision (ECCV)*, 2016: 75–91.
- [21] HE K, GKIOXARI G, DOLLAR P, et al. Mask R – CNN [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 42(2): 386–397.
- [22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *The IEEE Conference on Computer*

- Vision and Pattern Recognition(CVPR), 2016: 770 – 778.
- [23] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017: 936 – 994.
- [24] MNIH V, HEES N, GRAVES A, et al. Recurrent models of visual attention[C]//The Conference on Neural Information Processing Systems, 2014: 2204 – 2212.
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision(ECCV), 2018: 3 – 19.
- [26] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[C]//The Conference on Neural Information Processing Systems, 2015: 2017 – 2025.
- [27] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]//The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017: 764 – 773.
- [28] XU Z, XU X, WANG L, et al. Deformable convnet with aspect ratio constrained NMS for object detection in remote sensing imagery[J]. Remote Sensing, 2017, 9(12): 1312 – 1331.
- [29] 欧阳针,陈玮. 基于可变形卷积神经网络的图像分类研究[J]. 软件导刊, 2017, 16(6): 198 – 201.
- [30] 李旭青,张秦雪,安志远,等. 基于可变形全卷积神经网络的冬小麦自动解译研究[J/OL]. 农业机械学报, 2020, 51(9): 144 – 151.
LI Xuqing, ZAHNG Qinxue, AN Zhiyuan, et al. Automatic interpretation of winter wheat based on deformable full convolution neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 144 – 151. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200917&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.09.017. (in Chinese)
- [31] REN S, HE K, GIRSHICK R, et al. Faster R – CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137 – 1149.
- [32] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128(2): 336 – 359.
- ~~~~~
- (上接第 42 页)
- [17] 刘志忠,柳洪义,罗忠,等. 机器人工作空间求解的蒙特卡洛法改进[J/OL]. 农业机械学报, 2013, 44(1): 230 – 235.
LIU Zhizhong, LIU Hongyi, LUO Zhong, et al. Improvement on Monte Carlo method for robot workspace determination[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(1): 230 – 235. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20130143&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.01.043. (in Chinese)
- [18] 刘迎春,余跃庆,姜春福. 机器人可操作性研究进展[J]. 机械设计与研究, 2003(4): 34 – 37, 7.
LIU Yingchun, YU Yueqing, JIANG Chunfu. Survey of robot manipulability[J]. Machine Design & Research, 2003(4): 34 – 37, 7. (in Chinese)
- [19] 高秀兰,鲁开讲,王娟平. Delta 并联机构工作空间解析及尺度综合[J]. 农业机械学报, 2008, 39(5): 146 – 149, 145.
GAO Xiulan, LU Kaijiang, WANG Juanping. Workspace resolution and dimensional synthesis on Delta parallel mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(5): 146 – 149, 145. (in Chinese)
- [20] 陈士安,何仁,陆森林. 馈能型悬架综合性能评价体系[J]. 农业机械学报, 2006, 37(7): 14 – 18.
CHEN Shian, HE Ren, LU Senlin. Evaluating system of reclaiming energy suspension comprehensive performance[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(7): 14 – 18. (in Chinese)
- [21] TADOKOR S, KIMURA I, TAKAMORI T. A dexterity measure for trajectory planning and kinematic design of redundant manipulators[C]//Industrial Electronics Society, 1989. IECON'89. Conference of IEEE. IEEE, 1989: 415 – 420.
- [22] 王强,班义成,苏日娜,等. 农业机械适用性评价中性能试验法理论模型构建研究[J]. 内蒙古农业科技, 2013(6): 56 – 60, 63.
- [23] 刘东,龚方华,付强,等. 基于博弈论赋权的灌溉用水效率 GRA – TOPSIS 评价模型[J/OL]. 农业机械学报, 2017, 48(5): 218 – 226.
LIU Dong, GONG Fanghua, FU Qiang, et al. Evaluation model of irrigation water use efficiency based on game theory and GRA – TOPSIS[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(5): 218 – 226. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170527&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.05.027. (in Chinese)
- [24] 籍颖,张漫,刘刚,等. 农业机械导航系统综合评价方法[J]. 农业机械学报, 2010, 41(12): 160 – 164.
JI Ying, ZHANG Man, LIU Gang, et al. Synthetical evaluation of agriculture machine navigation system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(12): 160 – 164. (in Chinese)
- [25] 骆正清,杨善林. 层次分析法中几种标度的比较[J]. 系统工程理论与实践, 2004(9): 51 – 60.
LUO Zhengqing, YANG Shanlin. Comparative study on several scales in AHP[J]. Systems Engineering-Theory & Practice, 2004(9): 51 – 60. (in Chinese)
- [26] 程兴田,王祺,闫发旭. 农业机械适用性用户调查评价模型研究——基于模糊层次分析法[J]. 农机化研究, 2013, 35(7): 73 – 75, 80.
CHENG Xingtian, WANG Qi, YAN Faxu. Study on the evaluation model of agricultural machinery applicability by user survey with FAHP[J]. Journal of Agricultural Mechanization Research, 2013, 35(7): 73 – 75, 80. (in Chinese)