

胡志伟, 杨华, 娄甜田, 等. 基于全卷积网络的生猪轮廓提取[J]. 华南农业大学学报, 2018, 39(6): 111-119.

基于全卷积网络的生猪轮廓提取

胡志伟¹, 杨华¹, 娄甜田², 胡刚³, 谢倩倩³, 黄佳佳⁴

(1 山西农业大学 信息科学与工程学院, 山西 太谷 030801; 2 山西农业大学 经济管理学院, 山西 太谷 030801;

3 武汉大学 计算机学院, 湖北 武汉 430072; 4 南京审计大学 政府审计学院, 江苏 南京 211815)

摘要:【目的】实现猪舍场景下非接触、低成本生猪轮廓高效提取。【方法】以真实养殖环境下的生猪个体为研究对象, 提出一种基于 VGG16 与 UNET 相结合的全卷积神经网络模型 (VGG-UNET 模型)。该模型采用批处理方法, 迁移学习 VGG16 模型参数, 通过在模型中构建复制通道深度融合图像深层抽象特征与浅层特征, 实现对图像语义级别分割。在 30 头大白生猪的 1 815 张数据集上进行模型验证, 通过设置不同批大小对比试验, 并选取其中具有最佳效果的 3 组探讨批大小与评价指标值变化趋势间的关系。【结果】测试集上的对比试验结果表明, VGG-UNET 模型在像素精度与均交并比方面分别达到 94.32% 和 86.60%, 比单独采用 UNET 模型分别高出 0.89% 和 1.67%。不同指标值变化情况与批大小间的关系不尽相同。在本文试验环境下, 批大小对模型收敛速度的影响不明显。不同批大小条件下 PA 及 MIoU 指标值变化综合分析得出, VGG-UNET 模型具有较强稳定性和较高鲁棒性; 批大小为 8 的情况下 VGG-UNET 模型效果最佳。【结论】本文提出的生猪轮廓提取方法 (VGG-UNET 模型) 是有效的, 能实现精确、稳定的生猪轮廓提取, 且分割结果较为完整, 同时模型具有较高鲁棒性, 可为后续生猪个体识别研究提供参考。

关键词: 全卷积网络; 轮廓提取; 深度学习; 计算机视觉; 图像分割

中图分类号: S828

文献标志码: A

文章编号: 1001-411X(2018)06-0111-09

Extraction of pig contour based on fully convolutional networks

HU Zhiwei¹, YANG Hua¹, LOU Tiantian², HU Gang³, XIE Qianqian³, HUANG Jiajia⁴

(1 College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, China; 2 College of Economics and Management, Shanxi Agricultural University, Taigu 030801, China; 3 School of Computer Science,

Wuhan University, Wuhan 430072, China; 4 School of Government Audit,

Nanjing Audit University, Nanjing 211815, China)

Abstract: 【Objective】 To realize non-contacting and low-cost pig contour extraction under the piggery scene.

【Method】 We took individual pig in the real culture environment as the research object, and proposed a full convolutional neural network model based on the combination of VGG16 and UNET (VGG-UNET model). We adopted the batch processing method in this model to transfer and learn the parameters of VGG16 model. We achieved semantic level segmentation of the image by combining the deep abstract feature and shallow feature in depth via building the duplicate channel. The model was verified on 1 815 datasets of 30 Large White× Landrace pigs. Comparison experiments of different batch sizes were performed, and three groups with the best results were selected to explore the relationship between batch size and the evaluation index. 【Result】 Through comparison experiments on datasets, the pixel accuracy and mean intersection-over-union of VGG-UNET model

收稿日期: 2018-04-03 网络首发时间: 2018-10-19 11:34:35

网络首发地址: <http://kns.cnki.net/kcms/detail/44.1110.S.20181016.1407.032.html>

作者简介: 胡志伟 (1992—), 男, 硕士, E-mail: zhiweihu@whu.edu.cn; 通信作者: 杨华 (1973—), 男, 副教授, 博士, E-mail: yanghua@sxau.edu.cn

基金项目: 国家自然科学基金 (31671571)

were 94.32% and 86.60% respectively, which were 0.89% and 1.67% higher than those of the UNET model. The experiments showed different relationship between the change of different index values and batch size. Batch size had no obvious impact on the convergence rate of the model under this experimental environment. Through comprehensive analysis of PA and MIoU index values under different batch sizes, the VGG-UNET model showed the highest stability and robustness, and it was found to be the best when the batch size was 8.

【Conclusion】 The VGG-UNET model is effective for accurate and stable extraction of pig contour. Such segmentation result is relatively complete and the model has higher robustness, which can provide a reference for follow-up identification of individual pigs.

Key words: fully convolutional networks; contour extraction; deep learning; computer vision; image segmentation

随着深度学习的发展,基于机器视觉的视频分析技术逐步深入到牲畜养殖的各个领域^[1-4]。其中通过机器智能感知来识别生猪行为特征并构建养殖可追溯系统成为当今研究的热点,并已在生猪行为识别^[5]、体重监测^[6]、健康衡量^[7]、母猪发情预判^[8]以及分娩控制^[9]等领域取得不俗的研究成果。构建可追溯系统的核心步骤和经典难题是如何将生猪个体从猪舍场景中分离出来。生猪轮廓包含了猪体尺寸、猪蹄形状等身份信息,这些信息对区分生猪个体至关重要,因而可通过轮廓提取的方式分离猪舍环境下的生猪个体。当前机器视觉算法在人脸轮廓检测领域已经得到广泛应用^[10-12],因此,探讨如何基于图像处理技术进行高精度生猪个体轮廓提取以提高识别自动化程度和精度,并进一步提高计算机视觉算法在生猪行为分析研究领域的实用性,成为有待解决的问题。

基于机器视觉的生猪个体轮廓分割研究在诸多方面已经取得了较大进展。目前生猪分割算法主要分为静态图像分割、动态图像分割以及其他分割算法。常见的静态图像分割方法有分水岭分割、边缘检测分割、阈值分割等。如孙龙清等^[13]结合交互分水岭算法,提出基于改进 Graph cut 算法的生猪图像分割方法。朱伟兴等^[14]提出基于伪球边缘像素提取、低计算代价的猪只前景帧检测法,但该方法前景提取依据边缘点的概率,无法保证检测到前景的完整边缘像素点,所得的边缘轮廓并不适合做后继目标形状分析。动态图像分割方法有背景差分法、帧差法和光流法等,如:Guo 等^[15]基于复杂场景中群养生猪视频序列特征,提出一种基于 MoG 的背景差分法获取猪只目标区域,但其对处于前景与背景相差不大场景下的生猪个体轮廓分割效果并不理想;刘波等^[16]将生猪红外热图像和光学图像相结合,提出了一种基于非子采样轮廓波的图像融合算法,用于改善生猪异常检测中的前景轮廓提取;马丽等^[17]提出联立生猪外接矩形高宽比和低频傅里

叶系数构建猪只侧视图的特征向量的方法,能够从猪舍监控视频中自动分割猪只,但猪只外接矩形框对原始轮廓比较敏感,该研究未考虑卧姿的轮廓对理想侧俯视图的干扰。除此之外,随着卷积神经网络在人脸轮廓提取方面的成功运用^[18],探讨其用于动物个体轮廓分割领域成为可能。传统卷积神经网络(Convolutional neural network, CNN)由卷积层、池化层、全连接层以及 Softmax 分类器组成^[19],如常见的 AlexNet^[20]、VGG^[21]、GoogleNet^[22]等网络。图像经过一系列的卷积、池化及全连接操作后,映射成固定长度的特征向量,最终将其输入分类器中来识别图像所属类别。但这些网络模型在全连接层铺展过程中破坏了图像空间维度结构,丢失了大量灰度级信息,较难实现语义级的图像分割。而全卷积神经网络^[19](Fully convolutional networks, FCN)在卷积神经网络的基础上,通过反卷积层替代全连接层的方式对最终卷积层的 Feature map 进行上采样操作,有效保留图像空间维度信息,能够对图像进行灰度级分类,进而实现图像分割。杨阿庆等^[23]以 VGG16 为基础网络,设计哺乳母猪分割的 FCN,实现对猪舍场景下哺乳母猪的准确、快速分割。赵凯旋等^[24]结合帧间差值法与卷积神经网络,在提取出奶牛粗略轮廓的基础上实现奶牛个体身份识别。然而目前面向生猪个体的基于卷积神经网络的机器视觉技术,主要关注生猪个体在复杂环境下的提取问题,而对生猪个体轮廓提取问题研究甚少。

基于此,本文提出了一种基于计算机视觉算法的非接触、低成本生猪轮廓提取方法,在生猪真实养殖环境下,利用视频中图像序列样本量充足这一特点,基于人脸轮廓提取算法结合生猪个体特征,研究并提出了基于视频分析与全卷积神经网络的生猪个体轮廓提取模型。该模型通过结合 VGG16 与 UNET 网络构建,实现对猪舍环境下生猪轮廓精确提取,能有效分离视频帧中的生猪个体轮廓,避免猪舍、猪粪等噪声对后续识别效果的影响。

1 材料与方法

1.1 数据采集

本文 VGG-UNET 模型采用的数据集来自于 JDD-2017 京东金融全球数据探索者大赛提供的 30 个视频文件,每个文件对应猪舍场景下的 1 头长白生猪个体,时长均约为 1 min,在确保每张图中仅包含 1 头生猪个体的基础上,分别从每个视频中随机截取 5 张帧图,最终得到 30 头生猪共 150 张大小为 1280×720 像素的初始数据集图片。为减少模型训练占用内存,本文将 1280×720 像素图裁剪为 320×160 大小,考虑到裁剪前后图片宽高像素比不同的特点,对原图高度像素进行微量收缩,经该步处理前后生猪个体胖瘦虽发生细微变化,但并不影响后期模型试验效果。

如图 1 所示,为丰富数据集,对裁剪后的 150 张 320×160 像素的图片分别随机进行旋转、平移、调色度、亮度、遮光、缩放、加入椒盐噪声扰动等图像增强操作以扩充数据样例,其中旋转角度、亮度级别、遮光部位以及缩放比例等操作均随机选取对应数值以保证生成图片的随机性。经该步处理后,每张图片分别生成 24 张扩充样例,筛选其中 11 到 13 张变换效果较好的样例,最终得到 1 815 张图片,其中 30 头生猪各自包括 55~65 张图片。由于本文 VGG-UNET 模型仅用于提取生猪轮廓,并不关注猪只类别信息,因此即使在模型数据集中每头生猪对应图片数量不均衡也并不影响本试验效果。最终将扩充前后的数据作为本文模型输入数据集,并分为训练集 1 158 张,验证集 296 张以及测试集 361 张。以 1 头生猪帧图为例,对其进行数据增强操作后结果如图 1 图像增强部分所示。经旋转操作后,生猪个体在图片中所处空间位置发生明显变

化,这有利于模型对不同拍摄角度、不同空间位置的生猪轮廓有效抽取。经平移操作后,生猪个体部分部位移出图片空间区域,这有利于提升模型对不包含生猪某些部位轮廓提取精度,避免模型过分依靠生猪某些部位或某些特点进行轮廓提取。经色度调节、亮度、遮光以及加入椒盐噪声等图像增强操作,有利于模型提取光线较弱、包含生猪背影以及图片像素不高等复杂条件下的轮廓信息。对生猪个体进行缩放操作,有利于模型提取不同胖瘦状态下的生猪轮廓,避免生猪随生长周期导致体型变化而影响提取精度。

在得到模型输入数据集的基础上,考虑到全卷积神经网络训练是有监督的学习过程,因此需要手动标注模型的输出数据,即手动提取图片中的生猪轮廓部分用于模型输出。同时由于本文 VGG-UNET 模型仅用于提取生猪的轮廓部分,因此无需对生猪图片进行分类标签化处理,只需抠取出每张图中生猪部位即可。为此,本文对经上述图像增强操作得到的数据集进行下面 2 步预处理操作,以得到数据集的真实输出:

1) 利用 MAC 系统自带的 Pixelmator 软件抠取出经图像增强操作后的 1 815 张图片中的生猪个体部分,经抠取操作之后祛除了图片中的猪舍等背景信息,保留了图片中的生猪个体。该步操作如图 1 中生猪抠取部分所示,从中可看出,经该步操作后,生猪个体部位被有效抠取出来。

2) 为减少模型训练时间并保留猪体重要轮廓信息,本文对上一步处理获得的猪体部分像素值进行归一化操作。将三色通道值均除以 255,所得的值大于 0.5 则将该通道值替换为 1,所得值小于等于 0.5 则将该通道值替换成 0,最终可视化结果如图 1 归一化部分所示,从中可发现用于区分猪只个

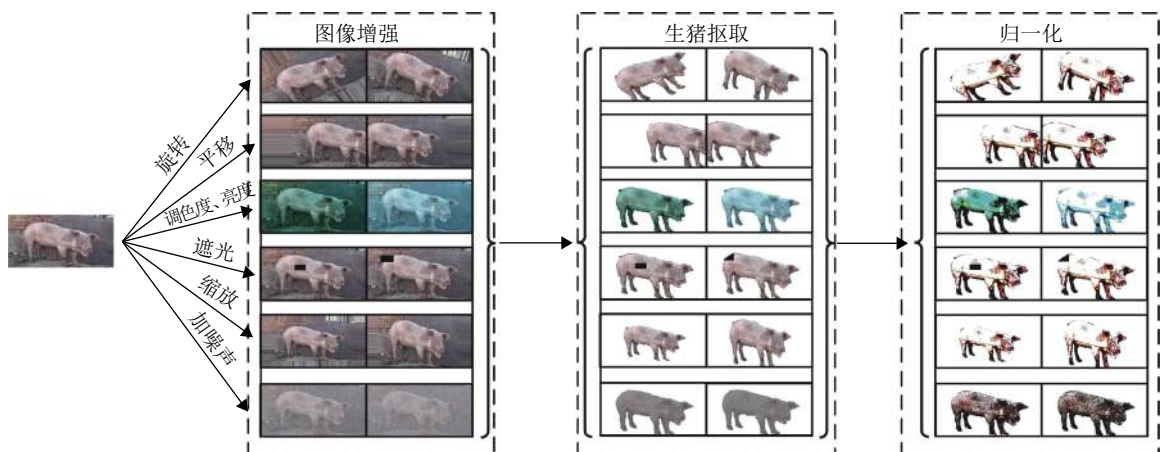


图 1 VGG-UNET 模型输入输出数据集创建过程

Fig. 1 The process of establishing input and output datasets for the VGG-UNET model

体信息与其他背景信息的猪耳、猪嘴及猪蹄等部位经该步处理后比较突出。

经上述预处理操作后, 整个数据集共包含 1 815 张 320×160 像素的图片, 将其用于模型输入以及 Pixelmator 软件抠取, 并经过归一化处理后得到 1 815 张仅包含生猪个体轮廓信息的图片, 将这些图片用于模型输出。此外上述预处理操作部分虽是在长白生猪数据集上进行的, 但在整个操作过程中并未刻意考虑生猪品种信息, 因此同样适用于诸如黑毛猪、宁乡花猪等品种的生猪个体预处理过程。

1.2 猪舍环境下生猪轮廓提取 VGG-UNET 模型

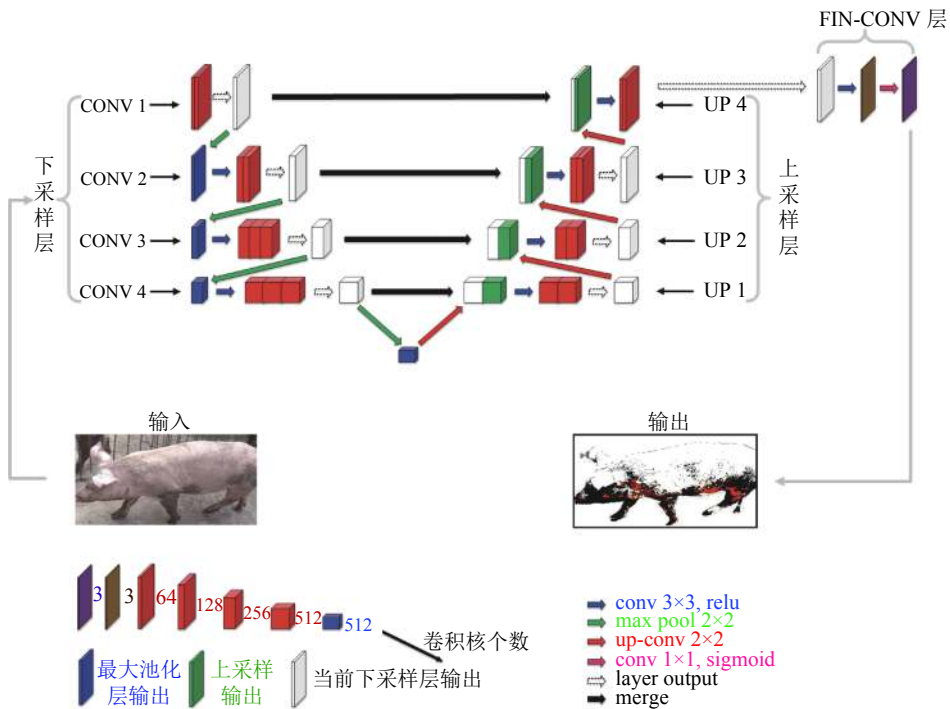
考虑到猪舍环境下的生猪图片帧中包含猪粪、猪食、猪栏等非生猪个体信息, 而这些信息在灰度级上存在较大差异, 因而可通过语义分割的方式将生猪轮廓从噪音环境中提取出来, 为此本文基于 VGG16 与 UNET^[25] 网络结构提出了 VGG-UNET 全卷积神经网络模型。

VGG-UNET 模型基于 VGG16 和 UNET 网络主要思想并对其进行了一定优化。VGG16 模型通过反复堆叠 3×3 的小型卷积核和 2×2 的最大池化层构建 16 层深的卷积神经网络, 被广泛用于提取图像特征。UNET 是基于 FCN 的语义分割网络, 共包括 23 个卷积层, 主要用于医学细胞领域的图像分割。本文通过迁移 VGG16 网络已训练好的模

型参数并结合 UNET 结构构建 VGG-UNET 混合模型, 充分利用 2 个模型的优势, 最终结构如图 2 所示。

该模型由编码层、解码层及 FIN-CONV 层(最终卷积层)3 个部分组成, 模型输入和输出图像均为 320×160 像素, 输入猪舍场景下的生猪个体图片, 输出为图片中的生猪轮廓部分信息。

从图 2 可知, 编码层由 VGG16^[21] 网络的前 10 层构成, 将这 10 层进一步划分为 4 层, 每层进行 2~3 次不等卷积操作, 卷积操作间通过引入 dropout 层以防止模型参数过多造成过拟合。每次卷积操作的卷积核大小均为 3×3, 编码部分层与层间通过 max-pooling 池化操作衔接, 池化窗口大小设置为 2×2。随着网络层数递增, 卷积核通道每经过 1 次池化操作均翻倍, 直到等于 512 为止。考虑到本文训练集较小的特点, 通过迁移学习 VGG16 已训练好的前 10 层模型参数初始化编码层卷积核参数, 在加速模型收敛的同时增强了模型的泛化能力。随着网络的深入, 编码层对输入的表达越来越抽象, 但具有更高的维度特征表示, 对图像的表达能力越来越强。经编码层处理后, 图像缩小为原图的 1/16。解码层基于 UNET 模型结构与 ResNet 网络思想构建, 共包括 4 层, 每层同样进行 2~3 次不等卷积操作, 卷积操作间设置 dropout 层, 卷积核大小均为



图中 conv 3×3, relu 中 conv 表示卷积操作, 3×3 表示卷积核大小, relu 表示激活函数; conv 1×1, sigmoid 中 conv 表示卷积操作, 1×1 表示卷积核大小, sigmoid 表示激活函数; up-conv 2×2 中 up-conv 表示反卷积操作, 2×2 表示反卷积核大小; max pool 2×2 中 maxpool 表示最大池化操作, 2×2 表示池化核大小; layer output 表示当前卷积层级输出; merge 表示融合与上采样层处于同一层的下采样层输出操作

图 2 VGG-UNET 网络结构图

Fig. 2 Diagram of the VGG-UNET network structure

3×3,随着解码层深入,卷积核通道数逐渐减小,这与编码层卷积通道数变化规律正好相反。层与层间通过 upsampling 操作衔接,对于 upsampling 操作,传统的全卷积神经将图像经过层层映射,只对最终映射结果进行上采样以此得到与输入大小相同的图像,在此过程中忽略了映射中间层的输出信息,而这些信息对于图像具有一定的表征能力。解码层借助 ResNet残差网络思想,通过构建复制通道的方式将映射中间层信息与对应解码层进行特征融合,使得浅层上下文信息能够传播到更高的分辨率层,将深层抽象特征与浅层特征组合起来以定位图像灰度级类别信息。FIN-CONV 层进行 2 次卷积操作,卷积核大小分别为 3×3 与 1×1,卷积核个数均为 3,经过在 FIN-CONV 层操作后,图像恢复到与输入相同的维度。

从图 2 中可看出,本文编码层即下采样部分共包括 4 层,从上至下分别表示 1~4 层,各层均采用 3×3 大小卷积核,第 1 层卷积核个数为 64,随着层数递增,每层卷积核个数呈现递增趋势,下一层卷积核个数是上一层卷积核个数的 2 倍,最终第 4 层卷积核个数增加为 512 个。本文解码层即上采样部分同样包括 4 层,整个网络模型从下至上分别表示 1~4 层,各层同样采用 3×3 大小卷积核,第 1 层卷积核个数为 512,随着层数递增,每层卷积核个数呈现递减趋势,下一层卷积核个数是上一层卷积核个数的一半,直至第 4 层卷积核个数减小为 64。经过编码层与解码层后,紧跟 FIN-CONV 层,分别执行 2 次个数为 3、卷积核大小为 3×3 卷积操作,提取出输入图片的轮廓部位。

VGG-UNET 通过构建复制通道的方式分层提取高低阶图像特性,保证最终输出图像中包含原图的上下文信息。该模型在下采样部分实现 4 层卷积操作,通过每层包含不同大小卷积核的方式分层提取图像中生猪个体部位信息,通过下采样操作可定位图片中的生猪个体部分,但无法获取其轮廓信息。随后的上采样部分实现 4 层反卷积操作,通过构建复制通道方式将映射中间层信息与对应解码层进行特征融合,在该过程中可逐步分层提取生猪轮廓,上层充分融合下层轮廓提取结果,最终得到完整生猪轮廓。

1.3 模型结果评价指标

采用图像分割领域公认的 4 种语义分割评价指标评估模型分割结果。评测指标分别是像素精度 (Pixel accuracy, PA)、均像素精度 (Mean pixel accuray, MPA)、均交并比 (Mean intersection over

union, MIoU) 以及频权交并比 (Frequency weighted intersection over union, FWIoU),各指标的定义如下^[26]:

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}, \quad (1)$$

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}}, \quad (2)$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}, \quad (3)$$

$$FWIoU = \frac{1}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ji} - p_{ii}}, \quad (4)$$

式中, p_{ij} 表示本属于类 i 但被预测为类 j 的像素数量, $k+1$ 表示类别数,本文 k 取 1。PA 表示标记正确的像素占总像素的比例,能够反映分割灰度级的准确性,而 MIoU 指标表示真实值与预测值像素间的交集和并集之比,能充分反映分割区域完整性和分割位置准确性,这 2 种指标常用于衡量图像语义分割的效果,本文生猪轮廓提取为图像语义分割中的子问题,因而本文将 PA 及 MIoU 值作为模型最终的评价指标分别用于评价模型对生猪图像中每个像素点的预测情况以及生猪轮廓部位提取完整程度。

1.4 试验参数设置

将数据集分为训练集、验证集及测试集 3 个部分,其中训练集大小为 1 158,验证集大小为 296,测试集大小为 361。为避免内存溢出,采取批训练方式对 UNET 和 VGG-UNET 模型在训练集和验证集上设置 6 组对比试验,6 组试验对应每个批次分别训练 4、6、8、10、12 和 14 张图片,即 6 组试验的训练批和验证批大小均分别设置为 4、6、8、10、12 和 14。遍历 1 次全部训练集数据称为 1 轮迭代,即 1 轮迭代会遍历所有训练集数据,而并不仅仅是 1 个批次数据,因此 1 轮迭代真正遍历的图片集大小为 1 158,以第 1 组试验为例,该组试验的批大小设置为 4,则 1 轮迭代需要执行 290 个批次。经过前期试探性测试发现,模型可能会在第 15 轮附近收敛,因而本文将迭代设置为 15×2 轮 (即 30 轮),每轮迭代完成后在测试集上计算“1.3”节中定义各个模型评价指标值。采用 keras 中对数损失函数 binary_crossentropy 作为代价函数。为解决反向传播过程中梯度消失和爆炸问题,引入批规范化,对网络隐藏层输入进行标准化。为加快模型收敛速度以获得最优模型参数,迁移学习已训练好的

VGG16 模型参数初始化 VGG-UNET模型下采样部分参数,并在迭代过程中 fine-tuning 微调参数。此外为防止过拟合,引入 keras 中的 EarlyStopping 机制,当训练集在连续迭代 5 轮 loss 值不再减小时停止训练。为提高调参效率,采用自适应矩阵估计算法 (Adaptive moment estimation, Adam)^[27]优化模型,初始学习率设置为 0.000 1。为保存最优模型参数,引入 keras 中的 ModelCheckpoint 机制,每轮迭代完成之后,通过观察验证集准确率是否发生提升决定是否保存当前模型。最终将保存的模型结构及参数用于后续生猪轮廓提取预测。

2 结果与分析

2.1 相同批大小下 VGG-UNET 与 UNET 模型对比

图 3 展示了 2 种模型在测试集上 6 组对比试验中每轮迭代完成后的 PA 指标值,每幅图对应一组对比试验,不同组试验间除了批大小选取不同外,其余试验参数设置完全相同。表 1 选取 2 种模型在测试集上 6 组对比试验中试验效果最佳的 3 组 (即批大小设置为 4、6、8) 计算其经 30 轮迭代后的平均评测指标 PA、MPA、MIoU 以及 FWIoU 值。

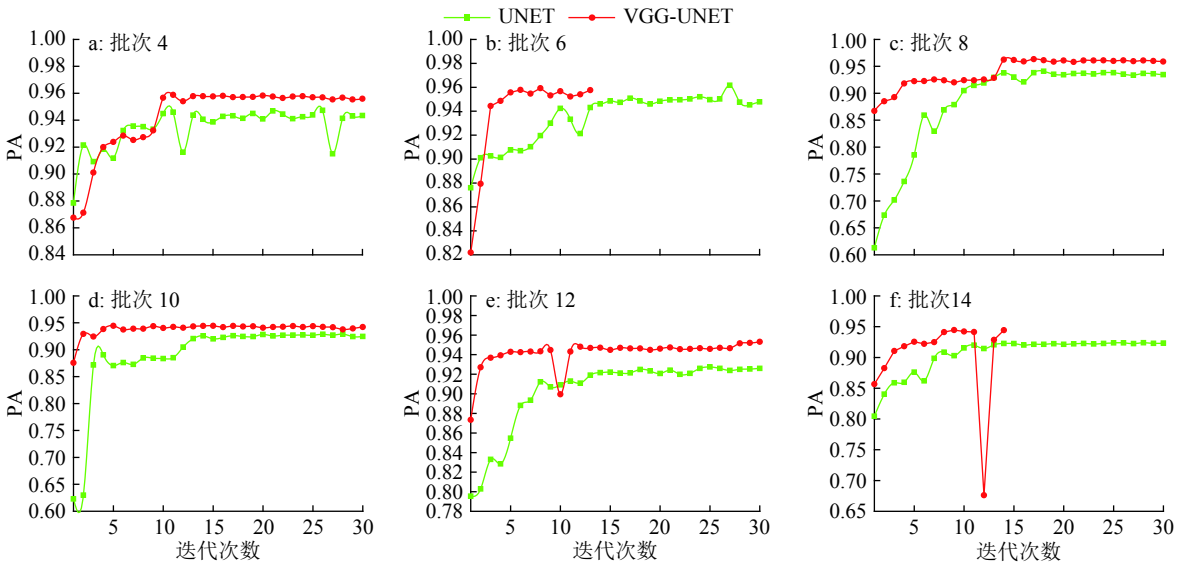


图 3 UNET 与 VGG-UNET 模型不同批大小下每轮 PA 变化

Fig. 3 The variations for each round of PA in different batch sizes of UNET and VGG-UNET models

如表 1 所示, VGG-UNET 模型 PA、MPA、MIoU 和 FWIoU 最优情况下分别达到 94.32%、88.51%、86.60% 和 86.65%, 比 UNET 模型最优情况分别高出 0.89%、1.71%、1.67% 和 0.58%, 且在反映分割区域完整性和分割位置准确性的 MIoU 指标上性能提升较大。可得出本文提出的 VGG-UNET 网络模型

明显优于 UNET 模型,这是由于 VGG-UNET 模型在下采样部分迁移已训练好的 VGG16 模型参数进行 fine-tuning 以得到最佳参数值,上采样部分借助 UNET 网络框架和 ResNet 残差思想,构建复制通道分层提取高低阶图像特征所致,说明 VGG-UNET 模型结构能够充分提取图像轮廓特征。

表 1 不同批大小下模型性能比较

Table 1 The performance comparison of models with different batch sizes

模型	批次	PA	MPA	MIoU	FWIoU
UNET	4	0.934 3	0.868 0	0.846 4	0.860 7
	6	0.932 7	0.863 8	0.849 3	0.858 1
	8	0.883 9	0.857 2	0.817 9	0.838 9
VGG-UNET	4	0.940 3	0.877 6	0.864 2	0.866 5
	6	0.941 8	0.874 1	0.864 9	0.864 1
	8	0.943 2	0.885 1	0.866 0	0.851 5

在 VGG-UNET 模型下,模型的衡量指标值变化趋势与批大小间存在一定关系。在 PA、MPA 和 MIoU 方面,批次 8 比批次 4 分别高出 0.29%、0.75% 和 0.18%, 比批次 6 分别高出 0.14%、1.10% 和 0.11%。但在 FWIoU 指标方面,批次 4 能达到最佳效果。说明批大小的选取对模型性能评价指标值变化的影响程度不同。对某些评价指标,批大小选取越大,指标值越高,但对于部分评价指标,指标值的高低与批大小选取间并不存在正比关系。这可能是由于指标计算的侧重点不同导致的。PA 指标更侧重像素级准确率,其值是标记正确的像素占总像素

的比例。MPA 作为 PA 的升级版, 计算的是所有类别上的均像素精度, 一般情况下, 这 2 个指标间的数值随着批大小的变化呈现相同的变化规律。但对于 MIoU 指标值, 更侧重的是计算分割区域完整性和分割位置准确性, 其值反映的是预测结果像素值与真实测定像素值间交集与并集的比值, FWIoU 指标作为 MIoU 指标的提升, 其变化规律理应与 MIoU 相近, 但该指标在 MIoU 指标的基础上加入权重参数, 而该权重参数的取值与预测出的图像轮廓占整个图像的比例有关, 因而可能会导致其变化规律与 MIoU 不一致的现象发生, 这也符合本文的试验结果。

从图 3 可探讨批大小对模型收敛速度的影响。对于同一种模型, 理论上批越大, 处理相同数据集所需迭代次数越少, 达到相同精度所需要的迭代次数将增多, 相应收敛速度将放缓。但在本试验环境下, 批大小对迭代收敛速度的影响程度并不是很大。VGG-UNET 模型在批大小为 6 的情况, 迭代至第 13 轮时触发了 EarlyStopping 机制, 提前结束了整个训练过程。而在批大小为 4、8 情况下, 分别在第 12、14 轮趋于收敛, 三者收敛速度相差无几, 出现这种情况可能是由于本文设置的批大小间相差不是很大所致, 但总体收敛需要的迭代次数会随着批大小选取增大而增加, 因而批大小对模型收敛速度的影响具有一定研究价值。

2.2 不同批大小条件下 VGG-UNET 模型试验结果分析

为纵向衡量批大小对 VGG-UNET 模型性能的影响, 图 4 绘制了其在测试集上不同批大小 (在此同样选取试验效果较佳情况即批大小为 4、6、8) 下每轮迭代的 PA 及 MIoU 指标值。

如图 4a 所示, 对于 VGG-UNET 模型, 不同批大小情况下每轮迭代过程中 PA 指标值变化呈现不同特点。批大小为 6 时, 迭代开始时 PA 较小, 但随

着迭代深入, PA 逐渐趋于稳定, 并在第 13 轮迭代时触发 EarlyStopping 机制, 即模型在第 13 轮前面的 5 轮训练中其 loss 值不再减小, 提前结束训练。而批大小为 4 或 8 时, 其 PA 在整个 30 轮训练过程中最初呈现较大波动现象, 但随着迭代的深入, 均维持在相对稳定的状态, 模型均能够在 30 轮迭代次数内收敛。说明 VGG-UNET 模型在不同批大小情况下均能保证较强稳定性和较高鲁棒性。如果将 PA 大小也考虑在内, 可发现, 从第 14 轮迭代开始, 批大小等于 8 的 PA 开始高于批大小为 4 时的数值, 因而在综合考虑模型稳定性与像素级准确率两方面, VGG-UNET 模型在批大小为 8 的情况下效果最佳。

如图 4b 所示, 对于 VGG-UNET 模型, 不同批大小情况下每轮迭代过程中 MIoU 指标值变化呈现不同特点。在批大小为 6 的条件下其值与 PA 呈现出相似的变化规律, 最初迭代开始时, MIoU 均小于批大小为 4 和 8 时的数值, 但随着迭代的深入, 其迅速在第 13 轮触发 EarlyStopping 退出机制而收敛。而批大小为 4 或 8 时, 随着迭代的深入, 虽未触发 EarlyStopping 机制, 但基本处于稳定状态, 并从第 15 轮迭代开始, 批大小等于 8 的 MIoU 虽然出现震荡情况, 但其值普遍高于批大小为 4 时的数值。综合考虑图 4a、4b 可发现, 对于 VGG-UNET 模型, 相同批大小条件下, 迭代过程中 PA 与 MIoU 呈现出一致的变化规律。原因在于 PA 与 MIoU 间虽侧重点不同, 但本身存在一定的关联, 两者结合更能反映模型的性能。

综合上述相同批大小条件下 VGG-UNET 与 UNET 2 种模型指标值对比以及不同批大小条件下 VGG-UNET 模型性能状况对比结果分析, 本文得出批大小为 8 情况下的 VGG-UNET 模型能够取得所有对比试验中最佳的效果, 因而该条件下的模型将作为后续模型中间层可视化与模型预测结果可视化分析使用的模型。

2.3 VGG-UNET 模型中间层可视化

为进一步探讨模型隐藏层学到的特征, 该部分以经“1.1”节试验数据预处理后得到的数据集中的 1 张生猪图片为例, 对 VGG-UNET 模型下采样部分的卷积层与上采样部分的反卷积层学习到的知识进行可视化展示, 最终结果如图 5 所示。

从图 5 可看出下采样部分的 CONV1 与 CONV2 层主要学到的是边缘轮廓、亮度等底层特征, CONV3 层学习到一些诸如纹理等有区别性的特征, 而 CONV4 层的特征具有更低的空间分辨率, 无法通过肉眼直

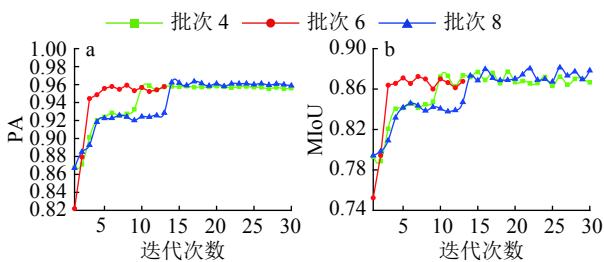


图 4 VGG-UNET 模型不同批大小下每轮 PA 及 MIoU 变化

Fig. 4 The variations for each round of PA and MIoU in different batch sizes of VGG-UNET model

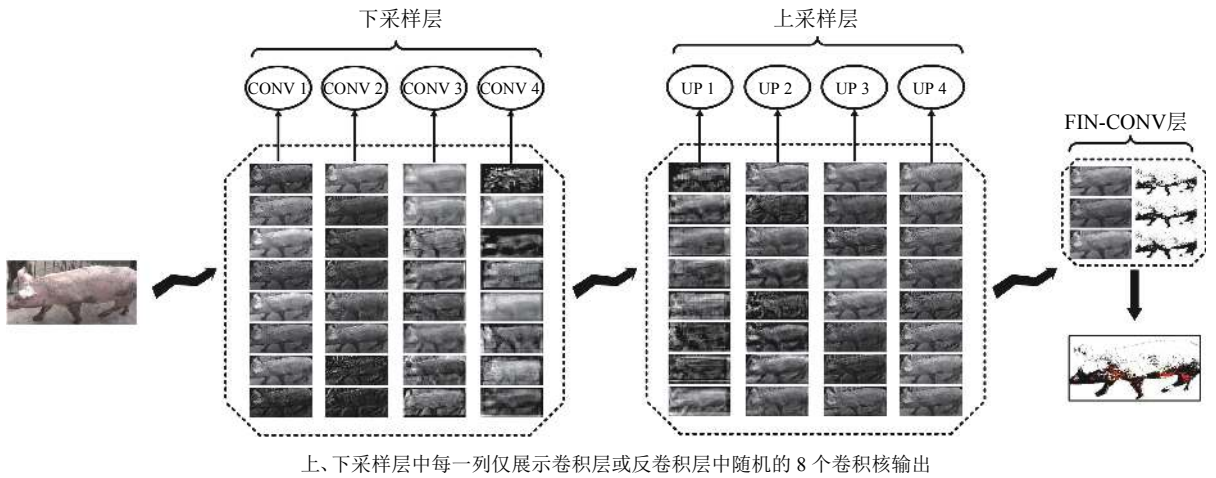


图5 VGG-UNET 模型中间层可视化

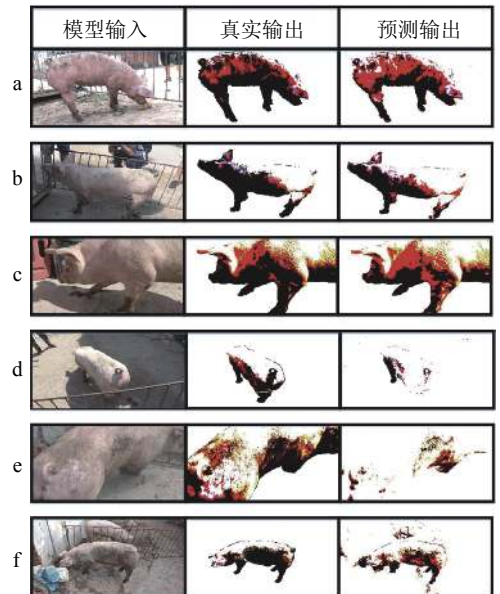
Fig. 5 Visualization of the middle layer of the VGG-UNET model

接判断出图片中是否存在生猪个体。即随着网络的深入,获取的图像越模糊,提取的特征越抽象,但区别于其他图像的特征越明显。对于上采样部分,UP1层获取的图像较为模糊,无法具有肉眼辨识性,经过4次上采样操作后,其表征结果逐渐明显。由此可得出上采样与下采样互为逆过程,下采样层数越深,特征抽取越抽象,而上采样层数越深,图像越清晰,特征抽取越明显。上述两部分的最终输出经FIN-CONV层3个 1×1 卷积核操作后,实现了对前述过程卷积输出的降维操作。通过对每个像素点在不同通道上进行线性组合,保留了图像的原有平面结构。最终将3个通道信息进行线性合并得到模型输出部分。

2.4 VGG-UNET 模型预测结果可视化分析

为进一步分析模型在不同输入情况下的预测效果,本文采用“2.2”节分析得出最佳模型,即批大小设置为8的VGG-UNET模型对测试集图片数据进行生猪轮廓提取,部分预测结果可视化后如图6所示。

图6中a、b和c组预测效果较好,d、e和f组预测效果较差。从图6a、6b和6c可看出,本文提出的模型对于侧身生猪轮廓提取效果较好,且模型能够排除人体、猪栏、光线强弱以及其他非生猪个体的干扰。对于背朝镜头或者不存在头部信息的生猪个体如图6d和6e所示,本文的提取效果不太理想,这可能是由于这2种情况下的生猪个体不包含能区别于生猪与其他事物的特点信息所致。对于粘连猪个体如图6f所示,VGG-UNET模型能够提取出其中较大的生猪个体轮廓,但对于个体较小的生猪提取效果不佳,这也是下一步研究需要解决的问题。



a-f组图片分别为不同图片的生猪轮廓提取结果;每组图片的第1列表示模型输入,第2列表示通过预处理后的真实输出,第3列表示批大小为8的VGG-UNET模型预测结果

图6 VGG-UNET 模型预测结果的可视化

Fig. 6 Visualization of prediction results from the VGG-UNET model

3 结论

本文提出一种基于全卷积神经网络的生猪轮廓提取VGG-UNET模型,以探讨具有更强实际应用能力的生猪轮廓提取方法。主要结论如下:

1) 为量化模型准确率,计算VGG-UNET模型在测试集上的各指标值。试验结果表明,对于PA与MIoU指标VGG-UNET模型比单独采用UNET模型最优条件下高出0.89%和1.67%。证明模型对于生猪轮廓提取具有较高准确率,可为实际生猪个体轮廓抽取提供参考。

2) 为探讨批大小与模型预测指标值变化趋势间的关系,选取在 VGG-UNET 模型上效果最佳的 3 组结果进行探究。结果表明,模型的衡量指标值变化趋势与批大小间存在一定关系,但不同指标值变化情况与批大小间的关系不尽相同。

3) 为探究批大小对模型收敛速度的影响,选取在 VGG-UNET 上试验效果最佳的 3 组结果进行分析。结果表明,在本文试验条件下,批大小对模型收敛速度影响不大。但批大小对模型收敛速度的研究具有一定价值,这也是下一步需要研究的内容之一。

4) 通过探究 VGG-UNET 模型不同批大小情况下每轮迭代过程中 PA 及 MIoU 指标值变化规律,得出该模型具有较强稳定性和较高鲁棒性,能够在提高准确率的同时保证模型收敛。并得出在批大小为 8 的情况下,本文提出的 VGG-UNET 模型效果最佳。

因此,本研究提出的进行生猪轮廓提取的方法是有效的,其提取的生猪个体区域细节完整,轮廓清晰,有较高稳定性和鲁棒性,可为下一步生猪个体身份识别提供有益参考。不足之处在于:一方面,本文模型对于背朝镜头生猪及粘连猪个体轮廓提取效果不太理想,另一方面本文针对批大小影响模型收敛速度方面研究欠妥,但这并不影响模型的实用性,在下一步研究工作中可适当深入探讨这两部分内容。

致谢: 由衷感谢 JDD-2017 京东金融全球数据探索者大赛为本文提供数据支持!

参考文献:

- [1] 杜晓冬,滕光辉,李卓,等.典型猪舍光环境下机器视觉图像量化评价及筛选[J].农业工程学报,2017,33(2): 213-219.
- [2] 蔡一欣,马丽,刘刚.奶牛隐性乳房炎便携式计算机视觉快速检测系统设计与试验[J].农业工程学报,2017,33(增刊1): 63-69.
- [3] 蔡骋,宋肖肖,何进荣.基于计算机视觉的牛脸轮廓提取算法及实现[J].农业工程学报,2017,33(11): 171-177.
- [4] 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].农业机械学报,2016,47(5): 231-244.
- [5] 肖德琴,冯爱晶,杨秋妹,等.基于视频追踪的猪只运动快速检测方法[J].农业机械学报,2016,47(10): 351-357.
- [6] 刘同海,李卓,滕光辉,等.基于RBF神经网络的种猪体重预测[J].农业机械学报,2013,44(8): 245-249.
- [7] 纪滨,朱伟兴,刘波,等.基于脊腹线波动的猪呼吸急促症状视频分析[J].农业工程学报,2011,27(1): 191-195.
- [8] 闫丽,沈明霞,谢秋菊,等.哺乳母猪高危动作识别方法研究[J].农业机械学报,2016,47(1): 266-272.
- [9] 刘龙申,沈明霞,柏广宇,等.基于机器视觉的母猪分娩

检测方法研究[J].农业机械学报,2014,45(3): 237-242.

- [10] CAO X, WEI Y, WEN F, et al. Face alignment by explicit shape regression[J]. Int J Comput Vision, 2012, 107(2): 2887-2894.
- [11] WU S, KAN M, HE Z, et al. Funnel-structured cascade for multi-view face detection with alignment-awareness[J]. Neurocomputing, 2017, 221: 138-145.
- [12] 张志刚,周明全,耿国华.人脸关键特征点自动标定研究[J].计算机工程与应用,2007,43(21): 197-198.
- [13] 孙龙清,李玥,邹远炳,等.基于改进 Graph Cut 算法的生猪图像分割方法[J].农业工程学报,2017,33(16): 196-202.
- [14] 朱伟兴,纪滨,秦锋.基于伪球算子边缘模型的猪前景帧检测[J].农业工程学报,2012,28(12): 189-194.
- [15] GUO Y, ZHU W, JIAO P, et al. Foreground detection of group-housed pigs based on the combination of mixture of Gaussians using prediction mechanism and threshold segmentation[J]. Biosyst Eng, 2014, 125(3): 98-104.
- [16] 刘波,朱伟兴,霍冠英.生猪轮廓红外与光学图像的融合算法[J].农业工程学报,2013,29(17): 113-120.
- [17] 马丽,纪滨,刘宏申,等.单只猪轮廓图的侧视图识别[J].农业工程学报,2013,29(10): 168-174.
- [18] 王林,董楠.基于 Gabor 特征与卷积神经网络的人体轮廓提取[J].南京理工大学学报,2018,42(1): 89-95.
- [19] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// IEEE. IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE, 2015: 3431-3440.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017, 60(6): 84-90.
- [21] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. Int J Comput Vision, 2015, 115(3): 211-252.
- [22] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE, 2015: 1-9.
- [23] 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.
- [24] 赵凯旋,何东健.基于卷积神经网络的奶牛个体身份识别方法[J].农业工程学报,2015,31(5): 181-187.
- [25] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// MICCAI. International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: MICCAI, 2015: 234-241.
- [26] GARCIA-GARCIA A, ORTS-ESCOLANO S, OPREA S, et al. A survey on deep learning techniques for image and video semantic segmentation[J]. Appl Soft Comput, 2018, 70: 41-65.
- [27] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-Image translation with conditional adversarial networks[C]// IEEE. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 5967-5976.