doi:10.6041/j.issn.1000-1298.2021.02.032

基于生成对抗网络与 ICNet 的羊骨架图像实时语义分割

赵世达[1]　王树才[1,2]　白　宇[1]　郝广钊[1]　涂本帅[1]

(1. 华中农业大学工学院, 武汉 430070; 2. 农业农村部长江中下游农业装备重点实验室, 武汉 430070)

摘要: 研究了羊骨架图像生成技术与基于 ICNet 的羊骨架图像实时语义分割方法。通过 DCGAN、SinGAN、BigGAN 3 种生成对抗网络生成图像效果对比,优选 BigGAN 作为羊骨架图像生成网络,扩充了羊骨架图像数据量。在此基础上,将生成图像与原始图像建立组合数据集,引入迁移学习训练 ICNet,并保存最优模型,获取该模型对羊骨架脊椎、肋部、颈部的分割精度、MIoU 以及单幅图像平均处理时间,并以此作为羊骨架图像语义分割效果的评判标准。结果表明,最优模型对羊骨架 3 部位分割精度和 MIoU 分别为 93.68%、96.37%、89.77% 和 85.85%、90.64%、75.77%,单幅图像平均处理时间为 87 ms。通过模拟不同光照条件下羊骨架图像来判断 ICNet 的泛化能力,通过与常用的 U-Net、DeepLabV3、PSPNet、Fast-SCNN 4 种图像语义分割模型进行对比来验证 ICNet 综合分割能力,通过对比中分辨率下不同分支权重的网络分割精度来寻求最优权值。结果表明,ICNet 与前 3 种模型的分割精度、MIoU 相差不大,但处理时间分别缩短了 72.98%、40.82%、88.86%;虽然 Fast-SCNN 单幅图像处理时间较 ICNet 缩短了 43.68%,但 MIoU 降低了 4.5 个百分点,且当中分辨率分支权重为 0.42 时,ICNet 分割精度达到最高。研究表明本文方法具有较高的分割精度、良好的实时性和一定的泛化能力,综合分割能力较优。

关键词: 羊骨架; 图像语义分割; 生成对抗网络; ICNet

中图分类号: S24; TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2021)02-0329-11　　**OSID**:

Real-time Semantic Segmentation of Sheep Skeleton Image Based on Generative Adversarial Network and ICNet

ZHAO Shida[1]　WANG Shucai[1,2]　BAI Yu[1]　HAO Guangzhao[1]　TU Benshuai[1]

(1. *College of Engineering, Huazhong Agricultural University, Wuhan* 430070, *China*

2. *Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan* 430070, *China*)

Abstract: Using computer vision technology to accurately and quickly identify the neck, ribs and spine of the sheep skeleton is the key to the development of a vision system for the slaughter and segmentation robot. To this end, the sheep skeleton image generation technology and the ICNet-based real-time semantic segmentation method of the sheep skeleton image were studied. DCGAN, SinGAN and BigGAN, three kinds of generation confrontation network were used to generate image effect comparison, BigGAN was selected to generate sheep skeleton image and expand the amount of sheep skeleton image data. On this basis, the generated image and the original image were combined to establish a combined data set, and then transfer learning was introduced to train ICNet and save the optimal model to obtain the segmentation accuracy of the model for the three parts of the sheep skeleton, MIoU and the average processing time of a single image, and it was taken as the criterion for the effect of semantic segmentation of sheep skeleton images, in the end, the optimal model's segmentation accuracy and MIoU for the three parts of the sheep skeleton were 93.68%, 96.37%, 89.77%, and 85.85%, 90.64%, 75.77%, and the average processing time for a single image was 87 ms. Then, the sheep skeleton images under different lighting conditions were simulated to judge the generalization ability of ICNet. Finally, comparing with the commonly used U-Net, DeepLabV3, PSPNet, Fast-SCNN four image semantic segmentation models to verify ICNet's comprehensive segmentation ability, and the optimal weight was

收稿日期: 2020-11-11　修回日期: 2020-12-07

基金项目: 国家重点研发计划项目(2018YFD0700804)

作者简介: 赵世达(1992—),男,博士生,主要从事家畜产品智能加工与检测研究,E-mail: zsd@webmail.hzau.edu.cn

通信作者: 王树才(1966—),男,教授,博士生导师,主要从事农产品智能化检测技术与装备研究,E-mail: wsc01@mail.hzau.edu.cn

found by comparing the network segmentation accuracy under different weights of the middle resolution branch. The test results showed that the segmentation accuracy and MIoU of ICNet and the first three models were not much different, but the processing time was reduced by 72.98%, 40.82% and 88.86%. In addition, Fast-SCNN single image processing time was 43.68% higher than that of ICNet, but the MIoU was reduced by 4.5 percentage points and the resolution branch weight was 0.42, ICNet segmentation accuracy reached the highest. Combined with the experimental results, it was showed that this method had high segmentation accuracy and good real-time performance, the comprehensive segmentation ability was the best, and it had a certain generalization ability. The above research provided theoretical support and technical support for the research and development of the intelligent segmentation robot vision system for sheep skeleton.

Key words: sheep skeleton; image semantic segmentation; generative adversarial network; ICNet

0 引言

羊肉是我国第四大肉类消费品,随着人民饮食结构的调整,其需求量逐年增加。据统计,2019 年我国羊肉总产量为 488 万 t,同比增长 2.6%[1]。在羊肉加工过程中,羊骨架分割是获得各食用部位的重要环节,与大众食用习惯息息相关。但是,由于缺乏智能分割设备,羊肉屠宰加工企业大多采用倒挂式半自动加工工艺,特别是羊骨架分割作业全部依赖人工完成,存在分割效率低、作业环境恶劣等问题,制约了我国家畜屠宰加工业的发展[2-3]。计算机视觉技术具有高效率、低成本、适应性好及稳定可靠的优点,已广泛应用于家畜养殖[4-5]、羊体尺测量[6]、肉质分级等[7-8]领域。采用计算机视觉技术可以准确、便捷地获取羊骨架各部位的自然特征表达,但基于深度学习的计算机视觉算法的有效性依赖于数据集的容量,需要海量的样本图像数据作为大数据驱动,才能取得较好的识别效果,这限制了其在小规模样本任务中的使用。因此,在保证图像语义信息的前提下,扩充羊骨架图像数据量、实现羊骨架图像的实时分割,从而满足生产线的实际需求,成为开发羊体分割智能机器人的关键。

近年来,生成对抗网络(Generative adversarial network,GAN[9])为计算机视觉技术的发展提供了新方向,其基于“零和博弈”和对抗训练的思想设计,可以有效实现高分辨率图像生成[10-11]与图像风格迁移[12],目前已被广泛应用于医学[13]、生物学[14]及农业领域[15]。文献[16]基于深度卷积生成对抗网络对植物图像识别方法展开研究,为农业图像数据不足问题提供了解决方案。文献[17]使用 GAN 网络生成蘑菇表型图像,解决了蘑菇表型数据长尾分布、非均衡性的问题。文献[18]基于生成对抗网络的 pix2pix 架构对人类肝脏 CT 图像进行风格迁移试验,结果表明,仿真数据与原始数据的结构相似性为 0.997 3。对于高分辨率的图像生成,一些学者发明了众多新型 GAN 网络架构,例如 C-GAN[19]、SinGAN[20]、BigGAN[21]、WGAN[22]等。如果采用高分辨率图像生成网络对羊骨架图像数据进行模拟,生成保留原始语义信息的羊骨架图像,则可以有效解决基于深度学习的计算机视觉方法对羊骨架图像进行实时语义分割中的样本不足问题。

羊骨架图像中羊颈部、肋部、脊椎 3 部位识别实质上属于图像语义分割问题,得益于卷积神经网络(Convolutional neural networks,CNN)对图像深、浅层特征强大的学习能力,研究者将卷积神经网络成功应用于各图像语义分割任务中。文献[4]采用卷积神经网络实现了奶牛个体身份的无接触精准识别,视频段识别率高达 93.3%。文献[23]针对群养模式下猪只体易粘连、难以分割的问题,提出了一种基于卷积神经网络的猪只图像分割方法。文献[24]通过 Faster-RCNN 对羊分娩场景下的初生羊羔进行了检测。文献[25]提出基于 FCN 哺乳母猪图像分割算法,使猪舍背景下的哺乳猪只被快速、精准分割。在羊体图像语义分割方面,目前相关研究较少。文献[26-27]采用 U 型卷积神经网络实现了羊肋排图像的准确分割,并基于 DeepLabV3+对羊骨架图像语义分割展开研究,最终取得较高的分割准确率,但该方法存在单幅图像处理耗时较长的问题。因此,考虑到羊骨架生产线对实时性的要求,实现羊骨架图像的快速语义分割非常必要。

本文首先基于对抗式生成网络对现有羊骨架图像数据集开展研究,生成具有完整语义信息的羊骨架图像,并通过 DCGAN[28]、SinGAN、BigGAN 3 种网络对比,优选出最佳 GAN 网络;然后,根据生成图像和原始图像建立组合数据集,设置不同亮度图像来模拟不同光照下的羊骨架图像,建立附加测试集;利用 ICNet[29]网络分割羊骨架图像,获取羊颈部、肋部、脊椎 3 部位的分割精度、MIoU 和单幅图像处理时间,与 4 种传统语义分割方法进行对比试验,并优化中分辨率分支权重,以期提高 ICNet 对羊骨架图

像语义分割精度。

1　材料与方法

1.1　试验材料与图像采集

试验样本选用经扯皮后去除头部、腿部、腹部的成年波尔山羊，样本图像采集于内蒙古自治区美洋洋食品股份有限公司的羊胴体分割生产线。选用华谷动力科技公司生产的 WP－UC600 型 CCD 相机，搭配 Z4S－LE－SV－1214H 型欧姆龙镜头，设置相机距地面 1.4 m，距悬挂状态的羊胴体样本0.8 m 拍摄图像，无特定背景与光源。为扩大样本之间的差异性，以保证后期模型训练的泛化能力，随机采集 6 批次羊胴体，每批次包含 200 个样本，共 1 200 幅尺寸为 3 024 像素 ×4 032 像素羊胴体图像。采集装置示意图如图 1 所示。

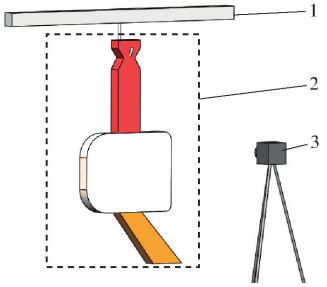


图 1　羊骨架图像采集装置示意图

Fig. 1　Illustration of image acquisition device

1. 输送带　2. 羊骨架　3. 相机

1.2　羊骨架图像生成对抗网络

1.2.1　生成对抗网络

生成对抗网络是由 GOODFELLOW 等[30] 于 2014 年基于零和博弈论提出的一种无监督学习算法框。GAN 网络一般由生成器和判别器两部分组成，生成器尽可能学习真实样本的数据分布，从而生成与真实数据逐渐相似的数据。判别器用以判定输入数据是否为真实数据，其输出一般是概率值，如概率值越大则表征输入数据是真实数据的可能性越大，反之即为生成数据。GAN 网络在训练过程中，生成器通过判别器的反馈机制不断更新参数，使得生成数据愈发服从真实数据分布，达到“欺骗”判别器的目的，而判别器也基于判断结果不断优化，提升区分生成数据与真实数据的能力。二者如此反复对抗，当判别器无法判断输入数据为生成数据或真实数据时，则认为 GAN 网络达到“纳什均衡”[31] 状态，模型训练达到最优。

GAN 网络结构如图 2 所示。随机噪声(z)一般为高斯变量，且生成器 G 对其维度没有限制。随机噪声输入生成器后，生成器生成数据 $G(z)$(生成图像)，进而输入至判别器 D，另外为保证生成器能够依据来源于判别器 D 的反馈误差进行参数更新，生成器必须可微。最后判别器 D 对真实数据(x，真实图像)与生成数据 $G(z)$ 做真假判定，并将误差传递至生成器，生成器再根据误差调整参数，目的在于使新生成“假”数据能够被判别器判定为“真”，同时，判别器根据判定结果更新自身参数以提高判别数据“真、假”的能力。

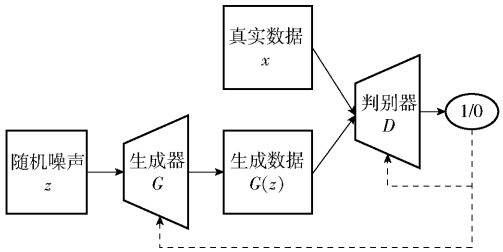


图 2　GAN 网络结构图

Fig. 2　Architecture of GAN

GAN 网络不断优化生成器与判别器的实质是将生成器极小化，判别器极大化，其目标函数为

$$\min(G)\max(D)V(D,G)=E_{x\sim P_{data(x)}}[\lg D(x)]+E_{Z\sim P_{Z(Z)}}[\lg(1-D(G(Z)))] \quad (1)$$

式中，$D(x)$ 表示当判别器的输入数据为真实数据时，得到判断结果为“真”的概率；$D(G(Z))$ 表示当输入数据为生成数据 $G(Z)$ 时，判别器判定为真实数据的概率；$P_{data(x)}$ 表示真实数据分布；$P_{Z(Z)}$ 表示生成数据分布；E 为期望值；生成器目的在于使生成数据无限接近真实数据，即 $D(G(Z))$ 趋近于 1，相应地，$V(D,G)$ 减小；判别器使 $D(x)$ 趋近于 1，而 $D(G(Z))$ 趋近于 0，$V(D,G)$ 增大；$\max(D)V(D,G)$ 表示判别器目标函数，使判别器预测概率的负对数期望取极大值；$\min(G)\max(D)V(D,G)$ 为生成器优化函数，使判别器的极大值最小化。

1.2.2　羊骨架图像生成网络

图像中目标的自然特征表达与图像分辨率息息相关，高分辨率图像较低分辨图像能够保留更全面的特征信息，且原始 GAN 网络存在生成样本缺乏多样性、模型训练容易塌陷、生成图像品质差等缺点，因此，考虑到上述问题及羊骨架不同部位特征的差异性，采用生成高分辨图像对抗网络很有必要。本研究首先将原始图像分辨率设定为512 像素 × 512 像素建立数据集，然后选用 DCGAN、SinGAN、BigGAN 3 种生成对抗网络生成羊骨架图像。

DCGAN 将卷积神经网络与 GAN 相结合，以提高生成图像质量及生成样本多样性。DCGAN 生成器 G 包含 3 个反卷积层(Deconvolution)和 3 个卷积层(Convolution)，判别器 D 包含 4 个卷积层及 1 个全连接层。DCGAN 网络结构如图 3 所示。

首先，尺寸为 64 × 64 × 128 的随机噪声矩阵输

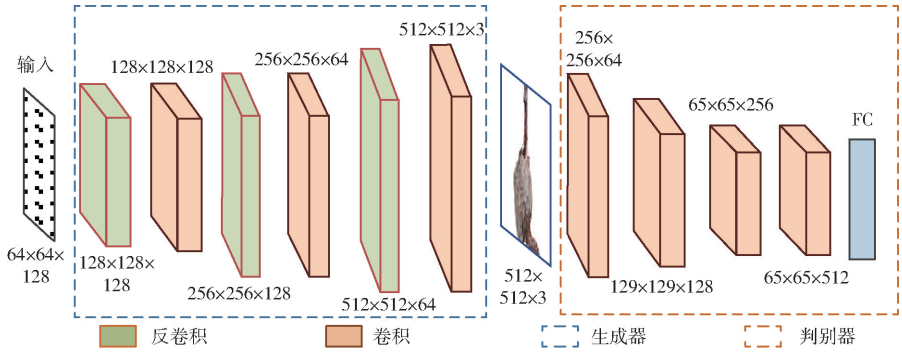


图 3 DCGAN 网络结构示意图

Fig. 3 Architecture of DCGAN network

入反卷积层;然后,输出尺寸为 128 × 128 × 128 的特征张量至卷积层,卷积核尺寸为 3 × 3,经批标准化(Batch normalization)和 ReLU 线性整流函数运算后输入下一层;最终,通过最后一层卷积操作经 Tanh 激活函数得到尺寸为 512 × 512 × 3 的羊骨架生成图像。判别器的输入为生成图像和原始图像,目的在于对图像的"真/假"进行判断。输入图像经过 4 层卷积运算,卷积核尺寸为 5 × 5,经标准化和 LeakyReLU 激活函数后输出至全连接层,最后全连接层经 Sigmoid 函数输出"1/0"。ReLU、Tanh 和 LeakyReLU 激活函数表达式参照文献[15]。

SinGAN 由金字塔结构的全卷积 GAN 网络组成,其主要特点在于能够基于单幅图像训练生成模型。SinGAN 通过学习不同尺度的图像数据分布,由金字塔结构的生成器 $\{G_0, G_1, \cdots, G_N D_N r^n \tilde{x}_N\}$ 依据多尺度的训练图像 $\{x_0, x_1, \cdots, x_N\}$ 由低分辨率向高分辨率逐步生成图像。SinGAN 图像生成过程如图 4 所示。

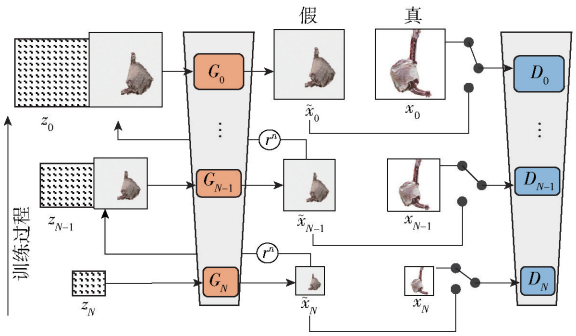


图 4 SinGAN 网络结构示意图

Fig. 4 Architecture of SinGAN network

随机噪声 z_N 输入底层生成器 G_N 生成图像 $\tilde{x}_N$,然后经采样因子 r^n 上采样至上一层与随机噪声 z_{N-1} 共同输入至第 $N-1$ 层生成器 G_{N-1},与此同时,生成图像 $\tilde{x}_N$ 与其尺寸相同的真实图像 x_N 传输至判别器 D_N,如此由下向上进行模型训练。SinGAN 这种模拟不同分辨率下的数据分布与低分辨图像上采样至高分辨图像生成网络策略,使其能够捕捉图像中细节特征。生成任意尺寸、多样性的高分辨率图像样本,同时保持与原始图像一致的语义特征。本研究于羊骨架图像原始训练集中随机选取单幅图像进行 SinGAN 网络训练。

BigGAN 网络于 2019 年提出,其通过扩大批尺寸(Batchsize)及增加网络参数量提升模型性能,并基于正交正则化与截断技巧有效地提升了生成图像质量及样本多样性,具备生成高像素、高品质图像的优点。本研究基于羊骨架图像数据集,通过调小 BigGAN 网络的 Batchsize 和生成器 G、判别器 D 学习率,使其能够部署在普通图形处理工作站上。

本研究根据以上 3 种羊骨架图像生成对抗网络的生成图像品质对比分析结果,获取最优生成网络。

1.3 羊骨架图像实时分割网络 ICNet

目前,应用于图像语义分割的深度学习模型多种多样,例如 DeepLabV3、PSPNet、U - Net、SegNet、FCN 等,上述模型在针对 Imagenet 数据集图像分割大赛中均取得较高的准确率,但是分割耗时较长,不能保证实时性。因此,轻量级语义分割模型 ENet 应运而生,但 ENet 保证实时性的同时放弃了一定的准确率,分割精度较低。文献[29]提出一种实时图像语义分割模型 ICNet,使模型兼顾实时性的同时保证了准确率。ICNet 使用 PSPNet 的金字塔池化模块融合多尺度上下文信息,并将网络结构划分为 3 个分支,分别为低分辨率、中分辨率和高分辨率,如图 5 所示。其中低分辨率分支将中分辨率输出的原图尺寸 1/16 的特征图进一步缩放至 1/32,之后采用空洞卷积扩大感受野,最终以原图尺寸 1/32 输出特征图,并且与中分辨率分支共享卷积参数与权重;中分辨率分支以原图 1/2 的分辨率作为输入,经卷积层后得到原图尺寸 1/16 的特征图,与低分辨率的输出特征图以 CFF 单元融合得到最终输出;高分辨率分支以原图作为输入,经卷积层后得到原图尺寸 1/8 的特征图,再通过 CFF 单元与中分辨率的输出特征图融合,经多倍上采样后将特征图扩充至原图尺寸。

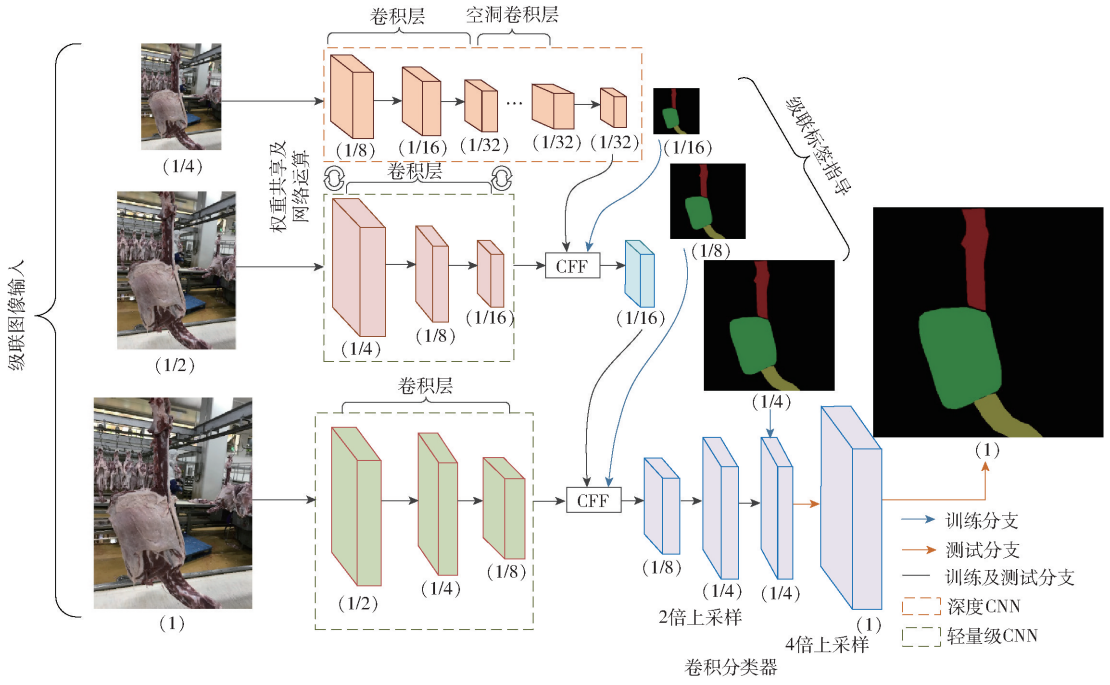


图 5　ICNet 网络 3 分支结构图

Fig. 5　Architecture of three branches for image cascade network

ICNet 利用低分辨率完成语义分割，高分辨率细化分割结果的策略提高了模型分割精度，此外，其采用的级联标签指导各分支的训练，加快了模型收敛与预测速度，提升了实时性。

ICNet 在每个分支训练中添加了损失权重，并对加权的 Softmax 交叉熵进行优化，其损失函数 $\mathscr{L}$ 可表示为

$$\mathscr{L}=\lambda_1\mathscr{L}_1+\lambda_2\mathscr{L}_2+\lambda_3\mathscr{L}_3 \tag{2}$$

式中　λ_1、λ_2、λ_3——低、中、高分辨分支权重

$\mathscr{L}_1$、$\mathscr{L}_2$、$\mathscr{L}_3$——低、中、高分辨分支损失

通常情况下，如果高分辨率分支权重 λ_3 设置为 1，则中分辨率和低分辨率分支的权重 λ_2 与 λ_1 分别为 0.4 和 0.16。

为量化分析 ICNet 模型针对羊胴体图像数据集语义分割的性能，本文引入图像语义分割任务中常用的像素精度 PA（Pixel accuracy）和平均交并比 MIoU（Mean intersection over union）作为模型分割性能的判断标准，将模型的分割结果与人工标注图像（真实标签）进行对比分析。PA 与 MIoU 计算式为

$$P_A=\frac{\sum_{i=0}^{N-1}n_{ii}}{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}n_{ij}}\times 100\% \tag{3}$$

式中　P_A——像素精度

N——语义类别数，取 4

n_{ii}——i 类语义的真实像素数量

n_{ij}——i 类语义被识别为 j 类的像素数量

$$M_{\mathrm{IoU}}=\frac{1}{N}\sum_{j=0}^{N-1}\frac{n_{ii}}{\sum_{j=0}^{N-1}n_{ij}+\sum_{j=0}^{N-1}n_{ji}-n_{ii}}\times 100\% \tag{4}$$

式中　M_{IoU}——平均交并比

n_{ji}——j 类语义被识别为 i 类的像素数量

上述度量标准中 MIoU 与模型的分割效果成正相关，因其简洁、代表性强，常作为图像语义分割模型性能评估的主要依据。

2　试验与结果分析

基于生成对抗网络及 ICNet 的羊骨架图像实时分割主要由 3 个步骤组成：①采集羊骨架图像，并归一化，建立羊骨架图像原始数据集。②基于生成对抗网络的羊骨架图像生成，获取最优模型。③针对“亮”、“暗”图像判断 ICNet 泛化能力；基于 ICNet 的羊骨架图像实时分割，并进行对比分析，最后进行 ICNet 优化。具体流程如图 6 所示。

2.1　试验平台

本文试验均基于 Pytorch 深度学习框架在 DELL T5810 型塔式图形处理工作站上完成，硬件环境为 Intel@ core64 至强 W－2145 CPU 3.70 GHz，RAM 为 64 GB，GPU 为 NVIDIA P4000－8GB，计算机系统为 Windows 10 专业版。

2.2　基于生成对抗网络的羊骨架图像生成

2.2.1　数据集

由于采集到的羊骨架图像分辨率较高，所占内存较大，如果直接作为后续模型训练的输入会显著增大模型的计算节点数，造成计算溢出，计算机无法

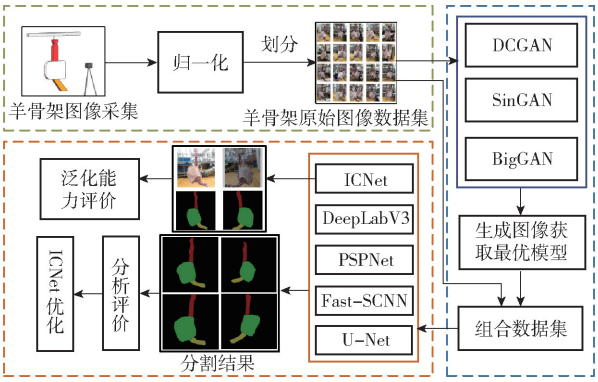


图 6 羊骨架图像生成及语义分割试验流程图

Fig. 6 Technical flowchart of experiment for sheep skeleton image generation and semantic segmentation

承担模型训练任务。因此,将原始羊骨架图像以比例不变性原则缩放至 512 像素 ×512 像素,经人工筛选,去除失真图像和增大目标与背景的对比度,最后选择 1 000 幅羊骨架图像作为原始数据集。数据集示例如图 7 所示。



图 7 羊骨架图像原始数据集示例

Fig. 7 Examples of original dataset of sheep skeleton images

2.2.2 生成对抗网络模型训练

生成对抗网络模型训练的实质在于以生成器 G 和判别器 D 反复对抗的方式更新网络参数,利用反向传播降低生成器和判别器的损失函数 G_loss与D_loss,使得生成图像接近原始图像的数据分布。

本文基于羊骨架图像原始数据集训练 DCGAN 与 BigGAN,随机选取原始数据集的某单幅图像训练 SinGAN。上述 3 种网络训练超参数设置如表 1 所示。

表 1 3 种网络超参数设置

Tab. 1 Hyperparameter settings of three network

超参数	DCGAN	SinGAN	BigGAN
优化器	Adam	Adam	Adam
生成器学习率	0.000 1	0.000 5	0.000 1
判别器学习率	0.000 1	0.000 5	0.000 4
动量参数 1	0.5	0.5	0.5
动量参数 2	0.9	0.9	0.9
Batchsize	16	1	1
迭代次数	30 000	30 000	30 000

2.2.3 羊骨架图像生成结果与分析

图 8 为 DCGAN、SinGAN、BigGAN 生成对抗网络针对羊骨架图像原始数据集训练过程中 G_loss 与 D_loss 随迭代次数的变化趋势。由图 8a、8b 可以看出,DCGAN 网络在训练初期 G_loss 随迭代次数迅速降低,当迭代 3 000 至 17 400 次时达到最低,随后振荡缓慢上升;其 D_loss 随迭代次数变化趋势与 G_loss 相似,在迭代初期迅速降低,然后振荡缓慢下降。SinGAN 网络 G_loss、D_loss 变化情况如图 8c、8d 所示,G_loss 在迭代初期迅速下降,之后于迭代 6 000 次左右显著上升随即快速下降,然后缓慢降低最终平滑;D_loss 呈现出大幅振荡状态,但总体小于 G_loss。如图 8e、8f 所示,BigGAN 网络 G_loss 在训

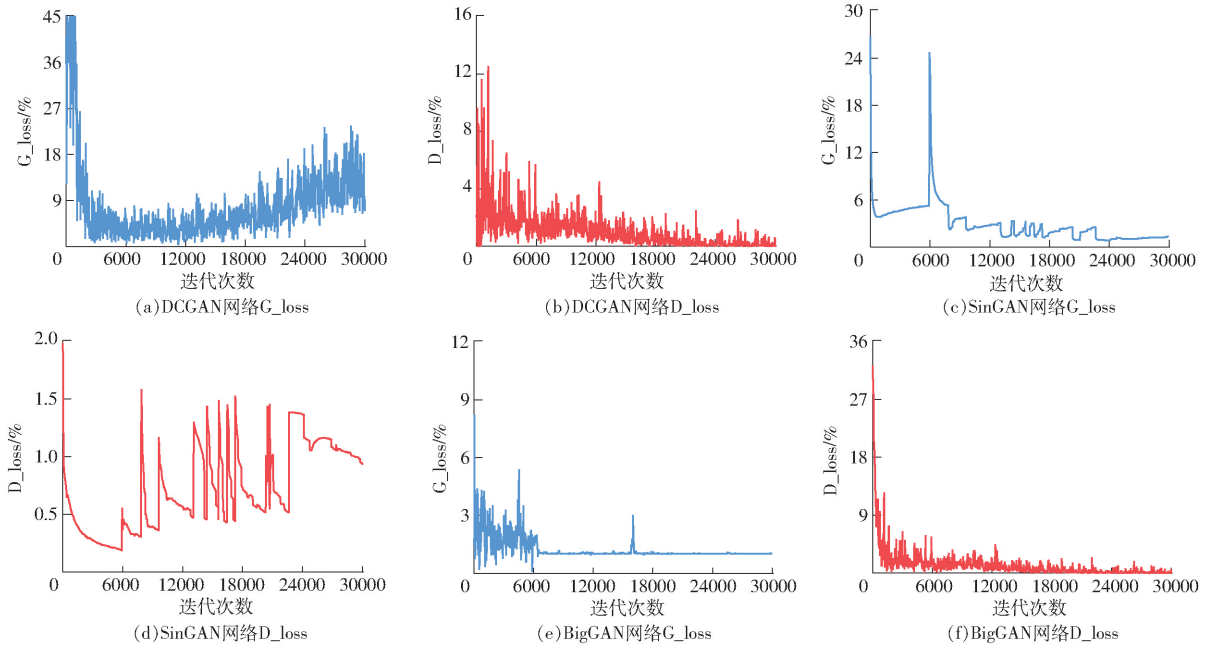


图 8 3 种对抗神经网络 G_loss 与 D_loss 随迭代次数的变化趋势

Fig. 8 Three types of GAN networks G_loss and D_loss change with number of iterations

练开始阶段以较大的幅值波动，然后趋于平稳，当迭代至 15 600 次附近时，迅速上升和下降，最终趋于稳定；D_loss 变化趋势与 DCGAN 网络相似，都随训练进度以小幅振荡形势缓慢降低。通过 3 种网络 G_loss和 D_loss 的对比变化情况可以得出，针对羊骨架图像数据集生成图像，判别器较生成器均能以少量优势取胜。

本文设置模型训练过程中每迭代 200 次生成一批羊骨架图像数据，3 种网络不同迭代次数生成的羊骨架图像如图 9 所示。

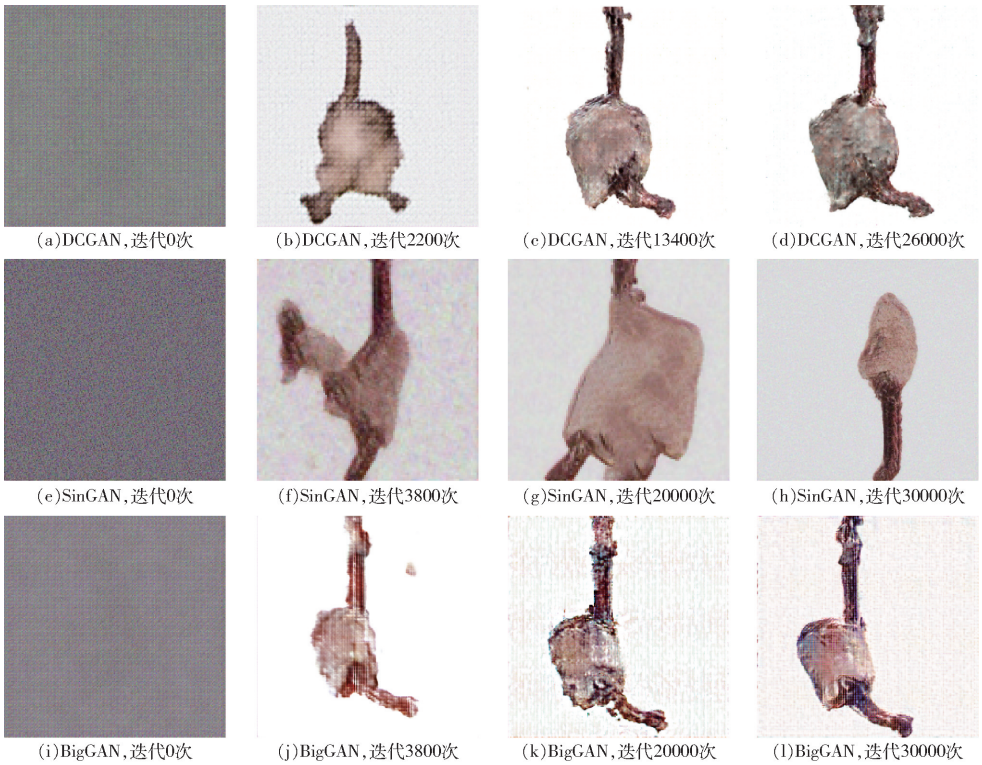


图 9　3 种对抗神经网络羊骨架图像生成结果

Fig. 9　Results of sheep skeleton image generation for three types of GAN networks

由图 9 可以看出，DCGAN 网络在迭代 13 400 次时，生成的羊骨架图像依据图像上下文信息，已经可以区分羊颈部、肋部和脊椎 3 部位，但特征细节与真实图像存在差距，其最终生成图像图 9d 品质优于图 9c；SinGAN 网络生成图像与真实图像相比特征差异明显，且出现失真情况，表明在本试验环境下，SinGAN 网络不适用于羊骨架图像生成任务；如图 9l 所示，随迭代次数的增加，BigGAN 网络最终生成的羊骨架图像中羊颈部、肋部、脊椎 3 部位语义信息保存完整，其颜色、纹理、轮廓特征明显，并且较图 9d 更为接近真实图像，表明针对本试验中羊骨架图像数据集，BigGAN 网络生成羊骨架图像性能优于 DCGAN，因此选用 BigGAN 作为最终的羊骨架图像生成网络。

2.3　基于 ICNet 的羊骨架图像分割

2.3.1　图像标注与数据集建立

本文羊骨架图像实时语义分割模型训练属于有监督学习。采集的羊骨架图像不包含标签及语义信息，需要人工进行图像标注，才能满足模型训练要求。基于 BigGAN 网络生成的图像与归一化后的原始图像组成羊骨架图像组合数据集，其中，训练集 6 000 幅，测试集 1 500 幅，验证集 200 幅。另外，为验证 ICNet 网络泛化能力，随机选取 100 幅同一节能灯光源条件下的羊骨架图像，并将其 RGB 色彩空间转换为 HSV，设置 1.5 倍和 0.8 倍两种亮度水平以模拟不同光照强度，最后再将“亮”、“暗”图像的色彩空间转换回 RGB，建立不同亮度的羊骨架图像附加测试集共计 200 幅，其中调亮、调暗图像各 100 幅。之后 2 种数据集均采用 Labelme 图像标注工具，参考 Cityscapes 数据集格式对羊颈部、肋部、脊椎 3 部位进行人工标注。

2.3.2　基于迁移学习的 ICNet 模型训练

迁移学习能够保证模型在面对小样本数据集时抑制过拟合现象，同时还可以加快模型收敛，提高泛化能力。因此，本文在 ICNet 模型训练过程中加载了基于 Cityscapes 数据集的预训练权重。采用 Adam(Adaptive moment estimation)优化器进行梯度下降，设置初始学习率为 0.001，Batchsize 为 4，迭代次数为 20 000。另外，设定模型在训练过程中自动保存最优模型，并将其作为羊骨架图像语义分割

的最终输入模型。ICNet 网络损失函数值随迭代次数的变化趋势如图 10 所示。由图 10 可知,损失值在训练初期迅速下降,当迭代次数为 2 000 左右时开始小幅度振荡缓慢下降,直至迭代次数为 16 000 次后损失值基本收敛于 0.032。

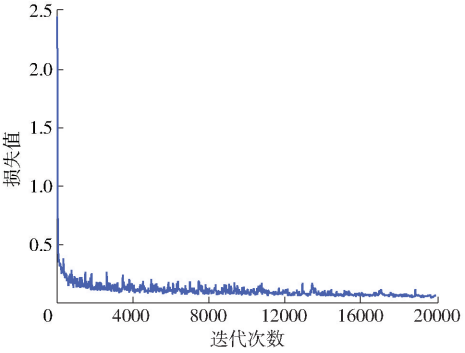


图 10 ICNet 训练过程中损失值随迭代次数的变化趋势

Fig. 10 Loss value changes with number of iterations during ICNet training

2.3.3 ICNet 对组合验证集的分割与结果分析

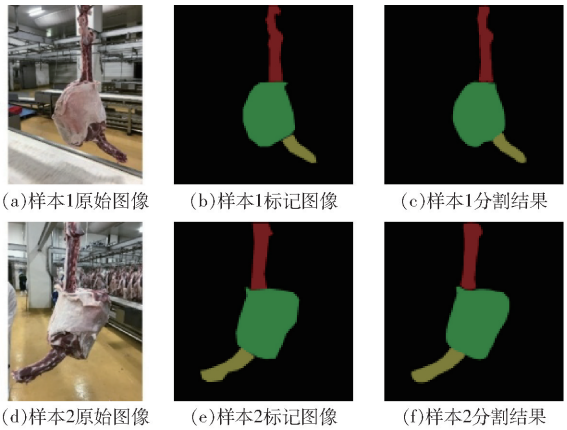
基于训练最优的 ICNet 模型展开测试,针对验证集获取羊体 3 部位的 MIoU 及准确率,以及模型的总体分割准确率和总体 MIoU。另外,为了判断模型是否具备良好的实时性,分别记录模型处理单幅图像时间,并求其均值。ICNet 针对羊骨架图像中脊椎、肋部、颈部 3 部位语义分割的准确率、MIoU 及单幅图像平均处理时间如表 2 所示,部分分割结果如图 11 所示。

表 2 基于 ICNet 的羊骨架图像分割结果

Tab. 2 Segmentation results of sheep skeleton image based on ICNet

类别	像素精度/%	MIoU/%	单幅图像平均处理时间/ms
脊椎	93.68	85.85	87
肋部	96.37	90.64	
颈部	89.77	75.77	
验证集	97.36	88.10	

根据图 11 和表 2 可知,羊骨架图像中羊颈部、肋部、脊椎 3 部位被准确分割,各部位区分明显,同时过分割与欠分割现象并不显著。特别是特征复杂的羊脊椎边缘识别明显,这可能是 ICNet 的低分辨率分支卷积层数更深,多层的卷积运算保证了细节抽象特征的提取。另外,图像背景中与羊骨架肋部颜色特征相似的羊骨架目标,并未对 ICNet 的分割结果造成干扰,其原因可能为仅去头、净膛后的羊胴体均保留了羊体四肢与带皮颈部,且未呈现脊椎,因此在尺寸、外形轮廓、区域纹理上与羊骨架特征存在明显差异。同时,ICNet 的多次上采样特征融合也



(a)样本1原始图像 (b)样本1标记图像 (c)样本1分割结果

(d)样本2原始图像 (e)样本2标记图像 (f)样本2分割结果

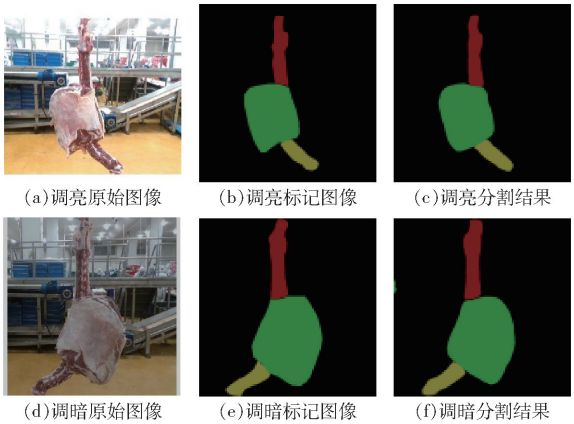
图 11 验证集部分分割结果

Fig. 11 Example of partial segmentation result of validation set

有利于提高模型的识别精度和降低过分割。最终针对组合数据集的验证集,ICNet 模型总体准确率与总体 MIoU 分别达到 97.36%、88.10%,单幅图像平均处理时间为 87 ms,表明 ICNet 能够实现羊骨架图像的准确语义分割,同时具备一定的实时性。

2.3.4 不同亮度下羊骨架图像语义分割结果

将基于组合图像数据集训练得到的最优 ICNet 模型针对不同亮度的羊骨架图像附加测试集展开泛化能力试验,其中部分分割结果如图 12 所示。



(a)调亮原始图像 (b)调亮标记图像 (c)调亮分割结果

(d)调暗原始图像 (e)调暗标记图像 (f)调暗分割结果

图 12 附加测试集部分分割结果示例

Fig. 12 Example of partial segmentation result of validation set

根据图 12 分割结果可以看出,羊骨架在“亮”、“暗”2 种光照强度情况下,ICNet 仍然可以实现其图像中 3 部位的准确分割,且脊椎-肋部、肋部-羊颈部黏连区域边缘区分明显,区域完整,轮廓清晰。最终,对羊骨架图像附加测试集中羊颈部、肋部、脊椎的分割精度与 MIoU 分别达到 92.53%、82.86%,94.37%、84.99%,89.06%、73.95%,说明 ICNet 能够以较高的精度克服因光源强度不同带来的识别干扰,具备较强的泛化能力。

2.3.5 与其他分割算法对比试验

目前,随着卷积神经网络的发展,用于图像语义

分割的深度学习方法越来越多，但是，面对不同任务、不同方法的分割性能表现差异明显。为进一步测试本文羊骨架图像实时语义分割方法的优劣性，引入目前图像语义分割任务中常用的 U－Net、DeepLabV3、PSPNet 和 Fast－SCNN 4 种图像语义分割模型与 ICNet 进行对比试验。上述 4 种模型的训练参数与 ICNet 相同，均基于自动保存最优模型的策略进行训练，之后针对验证集展开测试。部分分割结果如图 13 所示，图中从上至下依次为样本 1～4。

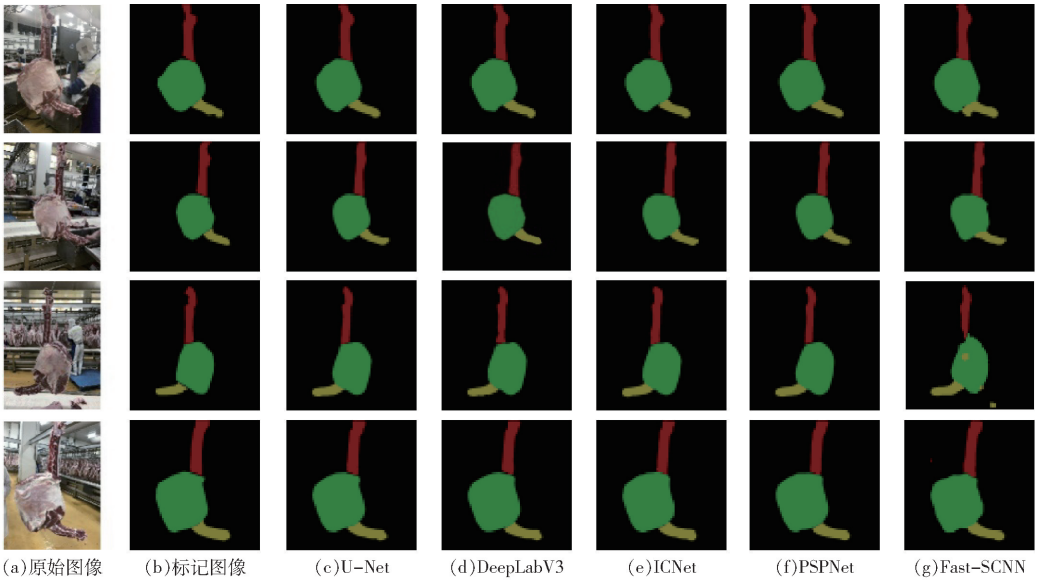


图 13　4 种图像语义分割模型与 ICNet 部分分割结果对比

Fig. 13　Comparison of four image semantic segmentation models and ICNet partial segmentation results

由图 13 可知，U－Net、DeepLabV3、ICNet、PSPNet 均可以实现 4 个羊骨架样本图像中颈部、肋部、脊椎 3 部位的准确分割，并且各部位边缘平滑，细节特征明显，能够满足实际生产中对切割精度的要求。但 Fast－SCNN 针对样本 1 和样本 3 的图像处理中出现过分割与欠分割现象，主要表现为将背景与肋部错误地分割为颈部区域，其原因可能是 Fast－SCNN 网络深度较浅，同时采用浅层学习下采样模块用于多分支低层特征的提取，在羊骨架图像数据规模有限的情况下，难以提取图像中的深层抽象特征用于网络学习，从而不利于后期的特征定位。而 U－Net、DeepLabV3、PSPNet 网络深度较深，都采用编码—解码结构，使得模型不仅可以提取更丰富的语义特征，还可以更好地恢复物体的边缘信息。另外，DeepLabV3 的空洞空间卷积池化金字塔结构及 ICNet 和 PSPNet 的金字塔池化模块使得模型具备获取更多上下文信息及多尺度特征的能力，保证了分割准确率。U－Net、DeepLabV3、PSPNet、Fast－SCNN 4 种模型总体准确率、MIoU 及单幅图像平均处理时间如表 3 所示。

由表 2 与表 3 可知，基于 U－Net 的羊骨架图像语义分割模型分割准确率与 MIoU 最高，达到 97.68% 与 88.56%，较 DeepLabV3、ICNet、PSPNet、Fast－SCNN 仅高 0.11、0.22 个百分点，0.32、0.46 个百分点，0.65、0.71 个百分点，1.22、4.96 个百分点，说明 5 种模型在分割精度方面相差不大，综合图 13 的可视化分割结果，Fast－SCNN 存在部分欠分割与过分割现象，因此只有 U－Net、DeepLabV3、ICNet、PSPNet 4 种模型可以满足羊骨架图像分割精度的要求。在分割实时性方面，U－Net、DeepLabV3、ICNet、PSPNet 单幅图像耗时分别为 322、147、87、781 ms，ICNet 用时最短，较 U－Net、DeepLabV3、PSPNet 分别缩短 72.98%、40.82%、88.86%，说明 ICNet 可以兼顾较高的分割准确率和良好的实时性，能够满足羊骨架切割生产线的实际需求。

表 3　4 种模型分割结果

Tab. 3　Segmentation results of four segmentation models

模型	像素精度/%	MIoU/%	单幅图像平均处理时间/ms
U－Net	97.68	88.56	322
DeepLabV3	97.57	88.34	147
PSPNet	97.03	87.85	781
Fast－SCNN	96.46	83.60	49

由于羊骨架颈部、肋部、脊椎 3 部位的特征差异，判断模型针对各部分的分割效果同样重要，本文获取了 U－Net、DeepLabV3、ICNet、PSPNet、Fast－SCNN 分别对羊骨架颈部、肋部、脊椎部位的分割精度与 MIoU，如图 14、15 所示。

由图 14、15 可以看出，ICNet 针对羊骨架 3 部位

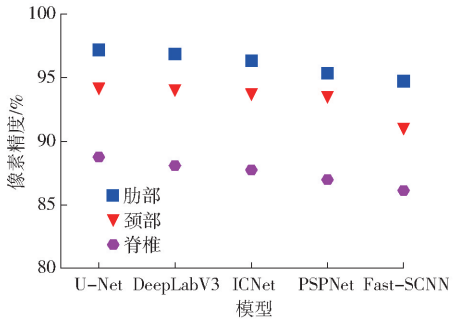


图 14 羊体 3 部位分割精度

Fig. 14 Accuracy of three parts of sheep skeleton

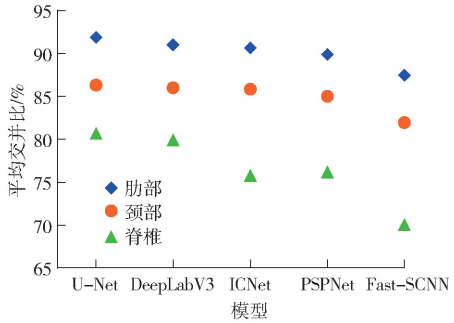


图 15 羊体 3 部位 MIoU

Fig. 15 MIoU of three parts of sheep skeleton

的分割准确率与 MIoU 均优于 PSPNet 和 Fast - SCNN,且与 U - Net 和 DeepLabV3 相差不大。最终,ICNet 对羊骨架图像中羊颈部、肋部、脊椎分割精度为 93.68%、96.37%、87.74%,MIoU 为 85.85%、90.64%、75.77%,能够满足生产线对羊体各部位分割精度的要求。另外,试验结果表明,ICNet 对图像中羊颈部的分割能力较弱,其原因可能为羊颈部和羊肋部存在黏连,且黏连区域特征与颈部区域特征非常相似,但黏连区域属于羊肋部,这种局部连接且特征明显相同的情况影响了模型对羊颈部区域的准确分割。

2.3.6 ICNet 优化对比试验

综合表 2、3 与图 14 可知,ICNet 分割精度比 U - Net、DeepLabV3 低 0.32、0.21 个百分点,且在脊椎部分分割效果相对较差。其原因可能是脊椎由等距相间的椎骨组成,表型复杂,特征多样,只有较高分辨率特征图才能尽可能保留该特征用于网络学习,而 ICNet 默认其高、中、低分辨率 3 分支权重分别为 1、0.4 和 0.16,在羊骨架图像语义分割任务中,该中分辨率分支权重较低,因此,通过调大中分辨率分支权重进行 ICNet 优化试验。过高的中分辨率权重可能会降低网络对高分辨率特征的敏感度,因而以 0.01 间隔扩大中分辨率权重至 0.46,共计 6 组对比试验,试验结果如表 4 所示。

表 4 ICNet 优化对比试验

Tab. 4 Optimizing ICNet comparative test

权重	像素精度/%	脊椎分割 MIoU/%
0.41	97.41	78.35
0.42	97.62	79.97
0.43	96.22	73.27
0.44	94.47	69.90
0.45	89.81	61.38
0.46	81.34	53.82

由表 4 可知,当中分辨率分支权重为 0.42 时,针对组合数据集 ICNet 的分割精度与分割脊椎部分的 MIoU 最高,达到 97.62% 和 79.97%。另外,原 U - Net、DeepLabV3、ICNet 对脊椎部分的 MIoU 为 80.67%、79.93%、75.77%,经对比,优化后的 ICNet 像素精度和脊椎分割 MIoU 与 U - Net 之间的差距分别缩小 81.25%、85.71%,比 DeepLabV3 和原 ICNet 提高 0.05、0.04 个百分点,0.26、4.2 个百分点。除此之外,随着权值的不断增大,网络精度显著下降,说明在本试验任务中,保持高、低分辨率分支权重不变情况下,设置中分辨率分支权重为 0.42 时,针对羊骨架组合数据集 ICNet 分割效果能够达到最优。

3 结论

(1)采用 BigGAN 网络生成的羊骨架图像能够有效保留羊颈部、肋部、脊椎 3 部位的语义信息,与原始图像最为接近,优于 DCGAN 与 SinGAN,能够满足基于计算机视觉技术的羊骨架图像分割数据集要求。

(2)针对组合数据集,ICNet 对羊骨架脊椎、肋部、颈部 3 部位的分割精度、MIoU 分别为 93.68%、96.37%、89.77% 和 85.85%、90.64%、75.77%,单幅图像处理时间达到 87 ms。对于不同亮度下羊骨架图像附加测试集,ICNet 依然能够实现羊颈部、肋部、脊椎的准确分割,说明本文方法在羊骨架图像语义分割任务中具有较高的分割精度、良好的实时性和一定的泛化能力。

(3)与 U - Net、DeepLabV3、PSPNet、Fast - SCNN 相比,ICNet 较前 3 种方法分割精度相差不大,但单幅图像处理时间分别缩短了 72.98%、40.82%、88.86%,虽然 Fast - SCNN 单幅图像处理时间缩短至 49 ms,但 MIoU 较 ICNet 下降了 4.5 个百分点,说明本文方法综合分割能力较优。另外,在本试验环境中,微调 ICNet 中分辨率分支权重为 0.42,能够将分割精度较原 ICNet 提高 0.26 个百分点。

参 考 文 献

[1] 国家统计局. 中华人民共和国 2019 年国民经济和社会发展统计公报[EB/OL]. (2020 - 02 - 28). http://www.stats.gov.cn/tjsj/zxfb/202002/t20200228_1728913.html.

[2] 丁存振,赵瑞莹. 我国肉羊屠宰加工业现状、问题及对策[J]. 肉类研究,2014,28(3):31 - 35.
DING Cunzhen, ZHAO Ruiying. Current status, issues and countermeasures of sheep slaughtering and processing industry in China[J]. Meat Research, 2014, 28(3): 31 - 35. (in Chinese)

[3] 方梦琳,张德权,张柏林,等. 我国羊肉加工业的现状及发展趋势[J]. 肉类研究,2008,22(1):3 - 7.
FANG Menglin, ZHANG Dequan, ZHANG Bolin, et al. The present situation and trends of the mutton processing in China[J]. Meat Research, 2008, 22(1): 3 - 7. (in Chinese)

[4] 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.
ZHAO Kaixuan, HE Dongjian. Recognition of individual dairy cattle based on convolutional neural networks[J]. Transactions of the CSAE, 2015, 31(5): 181 - 187. (in Chinese)

[5] 刘岩, 孙龙清, 罗冰, 等. 基于改进 CNN 的多目标生猪检测算法[J/OL]. 农业机械学报, 2019, 50(增刊): 283 - 289.
LIU Yan, SUN Longqing, LUO Bing, et al. Multi-target pigs detection algorithm based on improved CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp): 283 - 289. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=2019s044&flag=1. DOI:10.6041/j.issn.1000-1298.2019.S0.044. (in Chinese)

[6] 张丽娜, 杨建宁, 武佩, 等. 羊只形态参数无应激测量系统与试验[J/OL]. 农业机械学报, 2016, 47(11): 307 - 315.
ZHANG Li'na, YANG Jianning, WU Pei, et al. Design and experiment of non-stress measuring system for sheep's conformation parameters[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(11): 307 - 315. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20161142&flag=1. DOI:10.6041/j.issn.1000-1298.2016.11.042 (in Chinese)

[7] 范中建,朱荣光,张凡凡,等. 基于 BP 和 Adaboost - BP 神经网络的羊肉新鲜度高光谱定性分析[J]. 新疆农业科学, 2018, 55(1): 183 - 188.
FAN Zhongjian, ZHU Rongguang, ZHANG Fanfan, et al. Qualitative analysis of mutton freshness by hyperspectral imaging based on BP and Adaboost - BP neural network[J]. Xinjiang Agricultural Sciences,2018, 55(1): 183 - 188. (in Chinese)

[8] 朱荣光,姚雪东,段宏伟,等. 羊肉挥发性盐基氮的高光谱图像快速检测研究[J]. 光谱学与光谱分析,2016,36(3): 806 - 810.
ZHU Rongguang, YAO Xuedong, DUAN Chong, et al. Study on the rapid evaluation of total volatile basic nitrogen(TVB - N) of mutton by hyperspectral image technique[J]. Spectroscopy and Spectral Analysis, 2016, 36(3):806 - 810. (in Chinese)

[9] 林嘉豪,章宗长,姜冲,等. 基于生成对抗网络的模仿学习综述[J]. 计算机学报,2020,43(2):326 - 351.
LIN Jiahao, ZHANG Zongzhang, JIANG Chong, et al. A survey of imitation learning based on generative adversarial nets[J]. Chinese Joural of Computers, 2020,43(2): 326 - 351. (in Chinese)

[10] 周立君,刘宇,白璐,等. 一种基于 GAN 和自适应迁移学习的样本生成方法[J]. 应用光学,2020,41(1):120 - 126.
ZHOU Lijun, LIU Yu, BAI Lu, et al. Sample generation method based on GAN and adaptive transfer learning[J]. Journal of Applied Optics, 2020, 41(1): 120 - 126. (in Chinese)

[11] DENTON E, CHINTALA S, SZLAM A, et al. Deep generative image models using a laplacian pyramid of adversarial networks [J]. In NIPS, 2015, 2:1486 - 1494.

[12] LI C, WAND M. Precomputed real-time texture synthesis with Markovian generative adversarial networks[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016:702 - 716.

[13] CHOI E, BISWAL S, MALIN B, et al. Generating multi-label discrete electronic health records using generative adversarial network[EB/OL]. (2018 - 01 - 11)[2018 - 01 - 28]. https://arxiv.org/pdf/1703.06490v1.pdf.

[14] ZHENG R, LIU L, ZHANG S, et al. Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network[J]. Biomedical Optics Express, 2018,9(10):4863 - 4878.

[15] 李庆旭,王巧华,马美湖. 基于生成对抗网络的禽蛋图像数据生成研究[J]. 农业机械学报,2021,52(2):236 - 245.
LI Qingxu, WANG Qiaohua, MA Meihu. Generating poultry egg image data based on generative adversarial network [J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(2):236 - 245. (in Chinese)

[16] 张澎,崔梦天,谢琪,等. 基于深度卷积生成对抗网络的植物图像识别方法的研究[J]. 西南民族大学学报(自然科学版), 2019,45(2):185 - 191.
ZHANG Peng, CUI Mengtian, XIE Qi, et al. Research on plant image recognition method based on deep convolution generation adversarial networks[J]. Journal of Southwest Minzu University(Natural Science Edition), 2019,45(2):185 - 191. (in Chinese)

[17] 袁培森,吴茂盛,瞿肇裕,等. 基于 GAN 网络的蘑菇表型数据生成研究[J/OL]. 农业机械学报,2019,50(12): 231 - 239.
YUAN Peisen, WU Maosheng, QU Zhaoyu, et al. Mushroom phenotypic generation based on generative adversarial network [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(12): 231 - 239. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20191226&flag=1. DOI:10.6041/j.issn.1000-1298.2019.12.026. (in Chinese)

[18] 孟琭,钟建平,李楠. 基于 GAN 的医学图像仿真数据集生成算法[J]. 东北大学学报(自然科学版),2020, 41(3):332 - 336.
MENG Lu, ZHONG Jianping, LI Nan. Generating algorithm of medical image simulation data sets based on GAN[J]. Journal of Northeastern University (Natural Science), 2020, 41(3):332 - 336. (in Chinese)

[19] MIRZA M, OSINDERO S. Conditional generative adversarial nets[EB/OL]. [2019 - 03 - 16]. https://arxiv.org/pdf/1411.1784.pdf.

[20] SHAHAM T R, DEKEL T, MICHAELT T. Singan: learning a generative model from a single natural image[C]//Proceedings of the IEEE International Conference on Computer Vision,2019:4570 - 4580.

[21] ANDREW B, JEFF D, KAREN S. Large scale GAN training for high fidelity natural image synthesis[EB/OL]. (2018 - 9 - 28)[2019 - 2 - 25]. https://arxiv.org/pdf/1809.11096.pdf.

[22] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein GAN[EB/OL]. [2020 - 06 - 06]. https://arxiv.org/pdf/1701.07875.pdf.

[23] 高云, 郭继亮, 黎煊, 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? file_no=20190420&flag=1. DOI:10.6041/j.issn.1000-1298.2019.04.020. (in Chinese)

(下转第 380 页)

[14] 戴丰. 基于驾驶员起步意图的 AMT 控制方法研究[D]. 上海: 上海交通大学, 2012.
DAI Feng. Study on control of automated mechanical transmission based on driver's starting intentions[D]. Shanghai: Shanghai Jiao Tong University, 2012. (in Chinese)
[15] 陈海军, 赵治国, 王琪, 等. 干式 DCT 双离合器联合起步最优协调控制[J]. 机械工程学报, 2014, 50(22):150 - 164.
CHEN Haijun, ZHAO Zhiguo, WANG Qi, et al. Dry dual clutch transmission' optimal and coordinating launching control with two clutches based on the minimum principle[J]. Journal of Mechanical Engineering, 2014, 50(22):150 - 164. (in Chinese)
[16] ZHAO Zhiguo, HE Lu, YANG Yunyun, et al. Estimation of torque transmitted by clutch during shifting process for dry dual clutch transmission[J]. Mechanical Systems and Signal Processing, 2016,75:413 - 433.
[17] 杜岳峰, 傅生辉, 毛恩荣, 等. 农业机械智能化设计技术发展现状与展望[J/OL]. 农业机械学报, 2019, 50(9):1 - 17.
DU Yuefeng, FU Shenghui, MAO Enrong, et al. Development situation and prospects of intelligent design for agricultural machinery[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(9):1 - 17. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20190901&journal_id = jcsam. DOI: 10.6041/j.issn.1000-1298.2019.09.001. (in Chinese)
[18] 杨阳, 王蒙蒙, 刘永刚, 等. DCT 车辆起步数据驱动预测控制[J/OL]. 重庆大学学报, 2020.
YANG Yang, WANG Mengmeng, LIU Yonggang, et al. Data-driven predictive control for DCT vehicles starting process[J/OL]. Journal of Chongqing University, 2020. https://kns.cnki.net/kcms/detail/50.1044.N.20200610.1208.004.html. (in Chinese)
[19] GAO B, HONG J, QU T, et al. Linear-quadratic output regulator for systems with disturbance: application to vehicle launch control[C]//Youth Academic Conference of Chinese Association of Automation. IEEE, 2016:135 - 140.
[20] KIM K S, REW K H, KIM S. Disturbance observer for estimating higher order disturbances in time series expansion[J]. IEEE Transactions on Automatic Control, 2010, 55(8): 1905 - 1911.
[21] 刘蕾, 闫晓峰, 韩存武, 等. 基于 PI 观测器的奇异摄动系统故障诊断和最优容错控制[J]. 控制与决策, 2016, 31(10):1867 - 1872.
LIU Lei, YAN Xiaofeng, HAN Cunwu, et al. Fault diagnosis and optimal fault-tolerance control of singularly perturbed system based on PI observer[J]. Control and Decision, 2016, 31(10): 1867 - 1872. (in Chinese)
[22] KOENIG D, MAMMAR S. Design of proportional-integral observer for unknown input descriptor systems[J]. IEEE Transactions on Automatic Control, 2003, 47(12):2057 - 2062.
[23] 徐立友, 刘海亮, 周志立, 等. 拖拉机双离合器自动变速器换挡品质评价指标[J]. 农业工程学报, 2015, 31(8):48 - 53.
XU Liyou, LIU Hailiang, ZHOU Zhili, et al. Evaluation indexes of shifting quality for dual clutch transmission for tractor[J]. Transactions of the CSAE, 2015, 31(8):48 - 53. (in Chinese)

(上接第 339 页)

[24] 孙诗文. 智慧牧业中基于深度学习的羊分娩场景识别算法应用研究[D]. 鄂尔多斯:内蒙古工业大学,2019.
SUN Shiwen. Application research of sheep delivery scene recognition algorithm based on deep learning in wisdom animal husbandry[D]. Erdos: Inner Mongolia University of Technology, 2019. (in Chinese)
[25] 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)
[26] 赵世达,王树才,李振强,等. 基于 U 型卷积神经网络的羊肋排图像分割[J]. 食品与机械,2020, 36(9):116 - 121.
ZHAO Shida, WANG Shucai, LI Zhenqiang, et al. Image segmentation of shep ribs based on U-shaped convolutional neural network[J]. Food & Machinery, 2020, 36(9):116 - 121. (in Chinese)
[27] 李振强,王树才,赵世达,等. 基于机器视觉和机器学习的羊骨架自动分割方法[J]. 食品与机械,2020, 36(6):125 - 132.
LI Zhenqiang, WANG Shucai, ZHAO Shida, et al. The calculation methods of goat trunk's segmentation trajectory based on machine vision and machine learning[J]. Food & Machinery, 2020, 36(6):125 - 132. (in Chinese)
[28] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. [2020 - 06 - 06]. https://arxiv.org/pdf/1511.06434.pdf.
[29] ZHAO H, QI X, SHEN X, et al. ICNet for real-time semantic segmentation on high-resolution images[C]// European Conference on Computer Vision. Munich, Germany, 2017. arXiv: 1704.08545.
[30] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Advances in Neural Information Processing Systems, 2014: 2672 - 2680.
[31] RATLIFF L J, BURDEN S A, SASTRY S S. Characterization and computation of local nash equilibria in continuous games[C]// 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE, 2013:917 - 924.