

文章编号: 2095-2163(2022)12-0226-05

中图分类号: TP391.4

文献标志码: A

基于改进卷积神经网络的羊行为识别

李小迪, 王天一

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 本文根据羊不同行为的特征, 提出一种基于改进卷积神经网络的羊行为识别方法。构建卷积核尺寸全部为 3×3 的卷积神经网络(Convolutional Neural Networks, CNN); 使用缩放指数线性单元(scaled exponential linear units, *SeLU*)为激活函数, 使网络具有自归一化功能; 以最大池化(max pooling)为下采样; 在全连接层中采用丢弃(*Alpha dropout*)操作提高网络泛化能力, 使用余弦退火动态学习率进行动态微调; 进一步使用 *softmax* 分类器作为网络输出, 最终构建出羊行为识别网络模型。实验结果表明, 本文方法对羊进食行为识别准确率达到 90.30%, 站立行为识别准确率达到 94.16%。坐卧行为识别准确率能达到 91.90%。该模型能够实现羊不同行为的监测, 且有较高的准确性, 有助于提高畜牧管理效率和养殖智能化水平。

关键词: 深度学习; 羊行为识别; 图像识别; 卷积神经网络

Sheep behavior recognition based on improved convolutional neural network

LI Xiaodi, WANG Tianyi

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] According to the characteristics of different behaviors of the sheep, a sheep behavior recognition method based on improved convolutional neural network is proposed in this paper. Firstly, the paper constructs convolutional neural network (CNN) with convolutional kernel size of 3×3 , and using scaled exponential linear units (*SeLU*) as the activation function, the network has the function of self normalization. Then, max pooling is used as the down sampling. After that, in the full connection layer, the *Alpha dropout* operation is used to improve the generalization ability of the network, and the cosine annealing dynamic learning rate is used for dynamic fine-tuning. Finally, the *softmax* classifier is used as the network output, and the sheep behavior recognition network model is constructed. The experimental results show that the accuracy of this method for sheep eating behavior recognition is 90.30%, and the accuracy of standing behavior recognition is 94.16%. The recognition accuracy of sitting and lying behaviors reaches 91.90%. The model could monitor different behaviors of the sheep with high accuracy, which is helpful to improve the efficiency of animal husbandry management and the level of intelligent breeding.

[Key words] deep learning; sheep behavior recognition; image recognition; convolutional neural network

0 引言

近年来, 国内羊养殖行业发展迅猛, 羊养殖已然占据国内畜牧产业较大比重。为了满足市场需求, 羊养殖的规模不断扩大, 养殖方式从散养模式、牧场养殖向智能化、规模化的方式转变。智能化的养殖技术使得畜牧产业从注重羊养殖数量变为注重养殖质量, 经营方式从粗放式经营转变为工厂化、集约化。这样的转变不仅提高了养殖的水平, 更使得养殖产品的安全性得到改善, 同时也大大减少了由于不良养殖带来的环境污染, 提高了养殖的社会效益。此外, 智能化养殖减少了用工量, 改善了养殖场管理模式, 提升了养殖产品质量, 降低了养殖的用人成

本, 从而提高了养殖的经营效益^[1-5]。

羊的不同行为反映了羊个体的整体状况, 可以作为羊健康状况的重要参考。因此, 高效准确地识别羊的行为, 对羊的智能化养殖有重要意义。目前, 动物识别方法包括: 人工监测法、基于可穿戴设备的方法、基于深度学习的方法等。Marchioro 等人^[6]利用三轴加速度传感器采集不同行为的母猪数据集, 并通过分析不同行为的特征进行母猪行为识别, 实验证明, 该方法对母猪产仔行为识别的准确率最高达到 90%。Dutta 等人^[7]使用三轴磁力计和三轴加速度传感器穿戴在奶牛脖颈处, 记录奶牛的不同行为, 首先提取出奶牛不同行为的特征, 其次做归一化处理, 最后使用决策树、KNN 等不同算法对奶牛的

基金项目: 贵州省科学技术基金项目(黔科合基础(ZK[2021]一般304); 黔科合基础([2020]1Y254))。

作者简介: 李小迪(1996-), 女, 硕士研究生, 主要研究方向: 图像处理、计算机视觉; 王天一(1989-), 男, 博士, 副教授, 主要研究方向: 量子通信、图像处理、计算机视觉。

通讯作者: 王天一 Email: tywang@gzu.edu.cn

收稿日期: 2022-05-10

不同行为进行分类,实验结果表明,不同算法的识别准确率相近,其中决策树算法最优。郭冬冬^[8]使用三轴加速度传感器获取山羊不同行为数据集,然后使用k-means算法识别山羊不同行为,最终识别率达到 87%左右。何东健等人^[9]从监控视频中获取牦牛的行为信息,首先对牦牛进行目标检测,再使用k-means算法识别牦牛不同行为,实验证明,该方法识别准确率较高。

与人工监测法和可穿戴设备的方法相比,深度学习对于输入图像要求较低,对于有噪声的图像,无需对图像进行复杂的去噪处理,且能够从大量数据中自动地学习图像特征,利用卷积神经网络(Convolutional neural network, CNN)^[10-13]可以实现图像特征的自动提取,使用训练好的卷积神经网络进行行为识别,不仅提高了识别效率,也提升了识别的准确率。目前已成功应用于个体识别^[14]、图像分类^[15-16]、掌纹识别^[17-18]、行为识别^[19-22]等领域。

为了实现低成本,非接触式的羊行为识别,本文采集了大量养殖场环境下羊的行为视频,通过分析羊不同行为的特征,提出了基于改进卷积神经网络

的羊行为识别方法,构建卷积核尺寸为 3×3 的卷积神经网络,对比分析不同激活函数和不同丢弃率对模型性能的影响,使用余弦退火学习率对网络进行动态微调,分析有无遮挡情况下羊不同行为的识别准确率,实验结果显示,本文提出的方法识别速度较快,可达到 28 fps,对于羊的进食、站立、坐卧等行为识别准确率均达到 90%以上。

1 网络构建

1.1 卷积神经网络的构建

卷积神经网络是一种包含了卷积运算的人工神经网络,包含输入层、卷积层、池化层、输出层等部分。在羊行为识别的任务场景中,通过对羊不同行为视频逐帧分析发现:羊只的进食行为,进食位置在画面中随机出现,且羊在进食时羊头伸出羊圈,身体也不同程度地向下弯曲;羊只的坐卧行为,羊只腿部蜷缩,头部大多低下或紧贴地面,这 2 种行为都会导致羊只有不同程度的扭曲变形使得识别难度加大。基于此原因,构建适合羊只不同行为的卷积神经网络,模型结构设计如图 1 所示。对此拟做探讨分述如下。

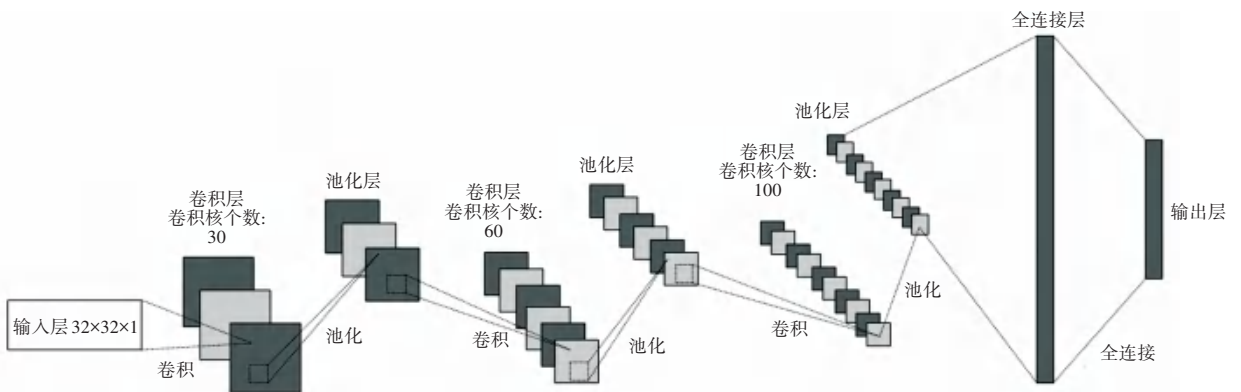


图 1 卷积神经网络结构图

Fig. 1 Structure of convolutional neural network

(1)输入层。由于养殖场中,羊的种类相同,羊的颜色和花纹都相近,且羊的外形、颜色并不影响羊行为的识别,因此输入的羊只不同行为的图像大小为 $32 \times 32 \times 1$,输入网络中进行网络训练。

(2)卷积层。羊的不同行为均有不同的特征,为了能够较准确地提取出羊不同行为的图像特征,本文构建了 3 个卷积层,分别为 C_1 、 C_2 、 C_3 。由于,2 个 3×3 卷积的感受野可代替一个 5×5 卷积的感受野,3 个 3×3 卷积的感受野可代替一个 7×7 卷积的

感受野,且 3×3 卷积网络的参数少,非线性更多,网络更深,使得模型的学习能力更强,且养殖场背景环境较为复杂,为了提取出更多的特征,本文构建的网络中卷积核的大小均为 3×3 。3 个卷积层的卷积核个数分别为:30 个、60 个、100 个。

(3)池化层。池化的过程也就是下采样的过程,在卷积神经网络中,需要通过池化层给得到的特征图降维,将特征进行压缩可有效缩小特征图的尺寸,从而减少计算量,防止网络的过拟合。本文采用

max-pooling 作为下采样的方法,选图像区域的最大值作为该区域池化后的值,池化层滤波器的尺寸为 2×2 。

(4)输出层。以全连接层和 *softmax*^[23] 层作为网络的输出层对网络提取的特征进行计算。全连接能够将输出的二维特征图转化为一维的特征向量,全连接层在卷积神经网络中可以作为“分类器”使用,本文在全连接层中加入丢弃操作,降低全连接层中各神经元之间的依赖性,以此提高网络的泛化能力。此外,使用 *softmax* 分类函数计算羊只不同行为的发生概率,公式如下:

$$\text{softmax}(f_x) = \frac{\exp(f_x)}{\sum_x \exp(f_x)} \quad (1)$$

通过对卷积神经网络主干层的分析,搭建出适合羊行为识别的网络,网络训练集分为进食行为、站立行为和坐卧行为三个类别。

1.2 SeLU 激活函数

在卷积神经网络中,激活函数是必不可少的,使用激活函数,使得网络上一层的输入与下一层的输出呈现非线性的关系,并且可以对任意函数进行逼近,使得网络具备更强大的表达能力。目前,大多数卷积神经网络都使用 *ReLU* 作为激活函数,但当 *ReLU* 激活函数的输入值小于 0 时,会造成梯度消失,使得部分神经元无法被激活。为解决这一问题,本文算法采用 *SeLU* 激活函数配合 *lecun-normal* 权重初始化方式。*SeLU* 激活函数的优点包括能够使神经网络自归一化、避免了梯度爆炸、梯度消失等问题、输出有正值和负值,能更好地控制均值等,其表达式如下:

$$\text{SeLU}(x) = \lambda \begin{cases} x & x > 0 \\ ae^x - \alpha & x \leq 0 \end{cases} \quad (2)$$

在文献[24]中已证明 $\lambda \approx 1.67$, $a \approx 1.05$ 。

1.3 丢弃操作

在模型训练过程中,由于羊的行为多样,使得少量羊的不同行为的图像较为相似,导致网络模型容易出现过拟合现象。因此,在训练网络时采用丢弃操作,让卷积神经网络的全连接层神经元按照一定的概率 p 随机置零,从而使得神经元之间的依赖性降低,提高模型的泛化能力。*Alpha dropout*^[24] 是丢弃操作中的一种,常规的丢弃操作会破坏网络的自归一化功能,而 *Alpha dropout* 以概率 p 将部分神经元置零时,保留下来的神经元会进行缩放和移位使得输出是零均值和单位方差,保证了数据的规范性,

其原理如图 2 所示。

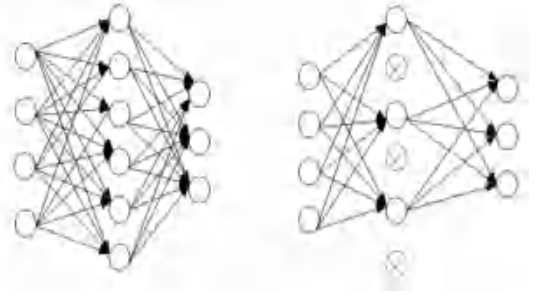


图 2 丢弃操作原理

Fig. 2 Alpha dropout operation principle

2 余弦退火动态学习率

在训练网络的过程中,学习率需要根据网络的训练情况进行动态调整。在网络训练初期,对于输入图像的像素信息完全未知,不恰当的学习率容易使模型陷入过拟合,在训练后期,过大学习率则会造成损失的大幅度震荡。因此,本文采用余弦退火动态学习率^[25-26],使学习率的变化符合余弦函数的变化方式,让模型轻松跳出局部最优解。其表达式如下:

$$\eta_i = \eta_{\min}^i + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i) \left(1 + \cos\left(\frac{T_{\text{cur}}}{T_i} \pi\right)\right) \quad (3)$$

其中, i 表示运行轮次; η_{\max}^i 表示学习率最大值; η_{\min}^i 表示学习率最小值,这 2 个值使得学习率在一定范围内衰减; T_i 表示第 i 次运行时总的 *epoch* 数; T_{cur} 表示当前执行了多少个 *epoch*。本文的初始学习率设置为 0.001,随着 *epoch* 的增加,学习率按照余弦规律减少,当训练到第 50 个 *epoch* 时,学习率下降速度达到最快,最终大小为 10^{-5} 。

3 实验结果分析

3.1 数据集的构建

实验视频在贵州省某羊只养殖场中采集,每段视频的平均时长为 3 min,格式为 *mp4*,帧宽度为 1 920 像素,帧高度为 1 080 像素,帧速率 50 帧/s。视频均在白天光线充足的条件下进行采集。

在预处理过程中,首先对采集的羊只行为视频进行人工筛选,筛选出进食行为视频 394 段,站立行为视频 312 段,坐卧行为视频 318 段。通过对关键帧的提取,手动删除关键帧中羊只肢体不全、遮挡严重、无羊只等无效图像,选择不同姿态、不同角度的羊只进食行为图像 3 500 张;羊只站立行为图像

2 800张,羊只坐卧行为图像3 100张。共计9 400张羊只不同行为图像构成实验所用数据集。在此基础上,按照 8 : 2 的方式划分训练集和测试集。

3.2 实验环境

本实验使用的操作系统为 Ubuntu18.04.4LTS, Python3.7, 深度学习框架: Tensorflow2.0, Graphics card: RTX 2080Ti , CPU: i7-9700, 总共 5 500 张图片,按照 8:2 的比例划分为训练集和测试集。

3.3 丢弃率对网络模型性能的影响

丢弃操作会使卷积神经网络的神经元按照一定的丢弃率 p 被置零,从而失去活性。当丢弃率 p 设置过高时,失活的神经元过多,容易导致模型欠学习,影响模型的识别准确率;当丢弃率 p 设置过低时,难以达到丢弃操作的效果。因此,本文将丢弃率分别设置为 0.2,0.3,0.4,0.5,0.6,0.7。不同丢弃率下网络的识别效果如图 3 所示。

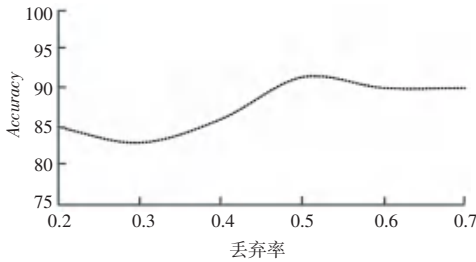


图 3 不同丢弃率下网络识别效果

Fig. 3 Network recognition effect under different discarding rates

由图 3 可知,当丢弃率 p 为 0.5 时,卷积神经网络模型准确率达到最高,因此,本文丢弃率设置为 0.5。

3.4 ReLU 激活函数和 SeLU 激活函数的比较

在卷积神经网络中,激活函数至关重要,本文使用 ReLU 激活函数的变体 SeLU 激活函数,实验结果如图 4 所示。由图 4 可知,在丢弃率设置为 0.5 的情况下,使用 SeLU 激活函数,模型的准确率由 89% 左右提升至 92% 左右,损失值 Loss 降低为 0.24 左右。

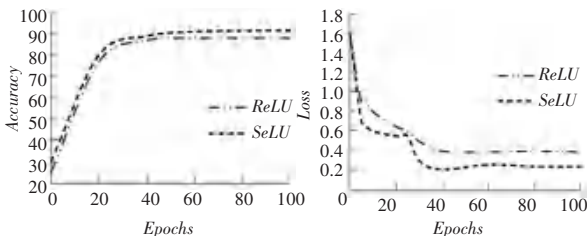


图 4 ReLU 激活函数和 SeLU 激活函数精确率和 Loss 的比较

Fig. 4 Comparison of Accuracy and Loss of ReLU activation function and SeLU activation function

3.5 有无遮挡情况下羊只行为识别效果

由于羊只在养殖场内活动,存在有、无栏杆遮挡的情况,如图 5、图 6 所示。因此本文对比有无遮挡情况下羊只行为识别结果,见表 1。



图 5 无遮挡羊进食、站立、坐卧行为

Fig. 5 Behaviors of eating, standing, sitting and lying of sheep without sheltering



图 6 有遮挡羊进食、站立、坐卧行为

Fig. 6 Behaviors of sheltering sheep from eating, standing, sitting and lying

表 1 有无遮挡情况下羊只行为识别效果

Tab. 1 Effect of sheep behaviors recognition with or without occlusion

行为分类	有遮挡	无遮挡	测试样本数/张	识别正确数/张	准确率 / %
羊进食行为	✓		300	259	86.33
羊进食行为		✓	350	328	93.71
羊站立行为	✓		250	229	91.60
羊站立行为		✓	350	336	96.00
羊坐卧行为	✓		300	268	89.33
羊坐卧行为		✓	330	311	94.24

由于 3×3 的卷积核能提取出较多的羊不同行为特征; SeLU 激活函数使得网络具有自归一化功能且避免了梯度爆炸和梯度消失的问题; Alpha dropout 降低了全连接层各神经元的依赖性,提高了网络的泛化能力;余弦退火动态学习率避免了网络出现过拟合现象,使得本文算法的识别率较高。由表 1 可知,在无遮挡情况下,羊的进食、站立、坐卧行为的识别率均达到 90% 以上,在有遮挡的情况下也能达到 85% 以上,有遮挡羊只行为识别的准确率稍低的原因是羊只行为部分遮挡严重,造成羊行为的特征发生扭曲变化,因此被网络识别错误。实验结果表明,本文方法能够完成羊只不同行为的识别,且识别结果较为可靠。

3.6 不同算法对羊只行为识别效果对比

表2为不同算法对羊只行为识别效果对比,表2中列举了3个典型的卷积神经网络对羊只行为的识别。由表2可知,ResNet18由于网络中存在残差结构,对单张图像的识别速度最快,达到31 fps。而本文提出的方法对单张图像识别的速度位居第二,达到28 fps。从识别精度方面考虑,本文提出的算法识别效果最优。基于以上实验结果可知,本文提出的方法能较好地完成羊只行为识别任务。

表2 不同算法下羊只行为识别效果

Tab. 2 Sheep behaviors recognition effect under different algorithms

算法	识别单张图片速度/fps	准确率/%
VGG16	25	84.43
ResNet18	31	89.69
LeNet5	26	90.38
本文算法	28	92.07

4 结束语

本文根据羊只进食、坐卧、站立等不同行为特征,提出了基于改进卷积神经网络的羊行为识别方法。通过构建卷积核尺寸为 3×3 的卷积神经网络,采用 $SeLU$ 激活函数增强网络的表达能力,通过设置丢弃率为0.5的 $Alpha$ dropout丢弃方式降低神经元之间的依赖性,并使用余弦退火学习率对网络进行动态微调。实验结果表明,在养殖场环境下,本文构建的卷积神经网络对羊不同行为的识别准确率均达到90%以上。该模型可以较高质量地完成羊不同行为的识别,有助于提高畜牧管理效率和养殖智能化水平。

参考文献

[1] 曹斌云,常宏军. 中国奶羊业的发展现状和战略[J]. 中国草食动物,2003(21):29-31.

[2] 史怀平,罗军,王建珏. 提高奶山羊养殖经济效益的几点思考[J]. 畜牧与兽医,2017,49(11):149-153.

[3] 王旺虎,郭子记,李锦旗. 智慧养殖在肉羊产业中的应用[J]. 畜牧兽医杂志,2021,40(04):4.

[4] 李卫卫,刘成忠. 羊养殖智能化系统研究[J]. 河南科技,2017(7):3.

[5] 刘良田. 2016. 浅析现代畜牧业发展现状及其对策[J]. 山西农经,2016(02):70.

[6] MARCHIORO G F, CORNOU C, KRISTENSEN A R, et al. Sows' activity classification device using acceleration data - A resource constrained approach [J]. Computers & Electronics in Agriculture, 2011,77(1):110-117.

[7] DUTTA R, SMITH D, RAWNLEY R, et al. Dynamic cattle

behavioural classification using supervised ensemble classifiers [J]. Computers and Electronics in Agriculture, 2015,111:18-28.

[8] 郭东东. 基于三轴加速度传感器的山羊特征行为识别研究[D]. 太原:太原理工大学,2015.

[9] 何东健,刘东,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(05):231-244.

[10] ACHOUR B, BELKADI M, FILALI I, et al. Image analysis for individual identification and feeding behaviour monitoring of dairy cows based on Convolutional Neural Networks (CNN) [J]. Biosystems Engineering, 2020,198:31-49.

[11] 周俊宇,赵艳明. 卷积神经网络在图像分类和目标检测应用综述[J]. 计算机工程与应用,2017,53(13):39-41.

[12] 段萌. 基于卷积神经网络的图像识别方法研究[D]. 郑州:郑州大学,2017.

[13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image Net classification with deep convolutional neural networks [J]. Communications of the ACM, 2017,60(6):84-90.

[14] 张俊龙,曾国荪,覃如符. 基于深度学习的海底观测视频中鱼类的识别方法[J]. 计算机应用,2019,39(02):376-381.

[15] 林泽文. 深度学习及其在图像分类识别中的运用探讨[J]. 信息与电脑,2018(12):146-147.

[16] 栾庆磊. 基于深度迁移学习的图像分类研究[J]. 黄山学院学报,2019,21(03):11-15.

[17] 郑仕伟,韩俊刚. 基于深度学习的非接触掌纹识别方法[J]. 微电子学与计算机,2018,35(04):98-102.

[18] 王海纶,李书杰,贾伟,等. 卷积神经网络在掌纹识别中的性能评估[J]. 中国图象图形报,2019,24(08):1231-1248.

[19] LIU Jingen, ALI S, SHAH M. Recognizing human actions using multiple features [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska: IEEE, 2008: 1-8.

[20] 赵歆. 基于 Res Net 网络的奶山羊行为识别方法研究[D]. 咸阳:西北农林科技大学,2020.

[21] DARRELL T, PENTLAND A. Space-time gestures [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE, 1993: 335-340.

[22] YAMATO J, OHYA J, ISHIIK. Recognizing human action in time-sequential images using hidden markov model [C]// Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Champaign, IL, USA: IEEE, 1992: 379-385.

[23] LIU Weiyang, WEN Yandong, YU Zhiding Yu, et al. Large-margin softmax loss for convolutional neural networks [C] // Proceedings of The 33rd International Conference on Machine Learning. New York: IMLS, 2016: 507-516.

[24] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks [C] //Advances in Neural Information Processing Systems (NIPS). Long Beach: NIPS Foundation, 2017: 971-980.

[25] LOSHCHELOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts [C]// International Conference on Learning Representations. Toulon, France:dblp, 2017:1-16.

[26] HE Tong, ZHANG Zhi, ZHANG Hang, et al. Bag of tricks for image classification with convolutional neural networks [J]. arXiv preprint arXiv:1812.01187, 2018.