

基于轻量级神经网络 MobileNetV3-Small 的鲈鱼摄食状态分类

朱 明^{1,3}, 张镇府^{1,2}, 黄 凰^{1,2*}, 陈燕燕¹, 刘亚东¹, 董 涛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;
3. 长江经济带大宗水生生物产业绿色发展教育部工程研究中心, 武汉 430070)

摘 要: 在集约化水产养殖过程中, 饲料投喂是控制养殖成本, 提高养殖效率的关键。室外环境复杂多变且难以控制, 适用于此环境的移动设备计算能力较弱, 通过识别鱼类摄食状态实现智能投喂仍存在困难。针对此种现象, 该研究选取了轻量级神经网络 MobileNetV3-Small 对鲈鱼摄食状态进行分类。通过水上摄像机采集水面鲈鱼进食图像, 根据鲈鱼进食规律选取每轮投喂后第 80~110 秒的图片建立数据集, 经训练后的 MobileNetV3-Small 网络模型在测试集的准确率达到 99.60%, 召回率为 99.40%, 精准率为 99.80%, F1 分数为 99.60%。通过与 ResNet-18, ShuffleNetV2 和 MobileNetV3-Large 深度学习模型相比, MobileNetV3-Small 模型的计算量最小为 582 M, 平均分类速率最大为 39.21 帧/s。与传统机器学习模型 KNN (K-Nearest Neighbors)、SVM (Support Vector Machine)、GBDT (Gradient Boosting Decision Tree) 和 Stacking 相比, MobileNetV3-Small 模型的综合准确率高出 12.74、23.85、3.60 和 2.78 个百分点。为进一步验证该模型有效性, 在室外真实养殖环境进行投喂试验。结果显示, 与人工投喂相比, 基于该分类模型决策的鲈鱼投喂方式的饵料系数为 1.42, 质量增加率为 5.56%。在室外真实养殖环境下, MobileNetV3-Small 模型对鲈鱼摄食状态有较好的分类效果, 基于该分类模型决策的鲈鱼投喂方式在一定程度上能够代替养殖人员进行决策, 为室外集约化养殖环境下的高效智能投喂提供了参考。

关键词: 水产养殖; 机器视觉; 图像识别; 深度学习; 神经网络; 鲈鱼

doi: 10.11975/j.issn.1002-6819.2021.19.019

中图分类号: S951.2

文献标志码: A

文章编号: 1002-6819(2021)-19-0165-08

朱明, 张镇府, 黄凰, 等. 基于轻量级神经网络 MobileNetV3-Small 的鲈鱼摄食状态分类[J]. 农业工程学报, 2021, 37(19): 165-172. doi: 10.11975/j.issn.1002-6819.2021.19.019 http://www.tcsae.org

Zhu Ming, Zhang Zhenfu, Huang Huang, et al. Classification of perch ingesting condition using lightweight neural network MobileNetV3-Small[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(19): 165-172. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.19.019 http://www.tcsae.org

0 引 言

在水产养殖过程中, 饲料是水产养殖中最主要的可变成本, 一般占养殖总投入的 40%~80%, 直接影响经济效益^[1]。目前中国淡水渔业养殖大部分采用人工投饵或机械投饵喂, 投喂量容易受到养殖人员的主观经验和意识的影响, 没有结合鱼群摄食过程中的参数信息, 难以做到合理投喂^[2]。如果投喂过量则会造成饵料的浪费, 增加了养殖的投入产出比, 另外过多的剩余饵料会污染水质造成水体的富营养化^[3], 投饵过少会减缓鱼的生长速度, 甚至引起鱼与鱼之间的攻击行为。如何在鱼类饵料投喂过程中做到适时适量, 是当前投喂面临的主要问题。

现有研究表明, 鱼群在摄食过程中会引起自身以及水面发生一系列变化, 因此探索鱼群摄食过程能够为合理投喂提供理论指导。Papadakis 等^[4]开发了一个可以进行远程监控鱼群行为的计算机视觉系统, 该系统可以快

速对鱼群特定行为进行分析。Pautsina 等^[5]利用近红外成像技术估测了鱼群的运动规律。乔峰等^[6]结合实时图像处理结果, 利用得到的鱼群摄食规律和投饵机的执行机构模型构建了智能投饵系统。郭俊^[7]以鱼群图像特征参数和声学特征参数作为输入量, 通过模糊处理和模糊逻辑推理, 构建以投饵量和投饵速度作为输出量的投饵策略, 实现了精准投饵。Atoum 等^[8]基于 SVM (Support Vector Machine) 分类器开发了一套适用于高密度水产养殖池的全自动投喂系统。陈明等^[9]通过 Relief 和 XGBoost 算法对鱼群摄食图像特征筛选和融合, 实现了镜鲤摄食强度评估。

随着深度学习的发展, 各类神经网络在视频图像处理任务中的效果越来越好, 在水产养殖中的应用也越来越多^[10]。Chen 等^[11]选取水温、溶解氧、平均鱼的质量和鱼的数量作为输入, 通过 MEA (Mind Evolutionary Algorithm) 优化 BP (Back Propagation) 神经网络, 建立了鱼类摄食量与环境因素的映射关系。Adegboye 等^[12]对鱼群活动向量进行离散傅里叶变换后得到傅立叶描述子, 输入到人工神经网络中实现了鱼群摄食量的预测。Måløy 等^[13]采用双流循环网络 (Dual-Stream Recurrent Network, DSRN) 对鲑鱼的摄食行为进行了预测。Villon 等^[14]通过卷积神经网络对水下鱼类进行了识别。赵德安等^[15]采用优化的 Retinex 算法和 YOLO V3 模型对水下河蟹进行了识别。Zhang 等^[16]采用多列卷积神经网络

收稿日期: 2021-06-11 修订日期: 2021-09-19

基金项目: 湖北省农业科技创新行动; 中央高校基本科研业务费专项资金资助项目 (107/11041910103); 中国工程院咨询项目 (2019-ZD-5)

作者简介: 朱明, 研究员, 博士生导师, 研究方向为农业工程管理。

Email: 13801392760@163.com

*通信作者: 黄凰, 博士, 讲师, 研究方向为农业机械化与农业智能化管理。

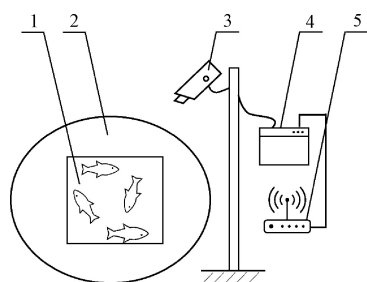
Email: wmyhuang@qq.com

(Multi-column Convolution Neural Network, MCNN) 作为前端和扩张卷积神经网络 (Dilated Convolution Neural Network, DCNN) 作为后端的方式, 建立了一种混合神经网络以实现鱼类种群自动计数。Tseng 等^[17]以卷积神经网络为基础, 提出一种对鱼类各个部位进行分类的模型。Zhou 等^[18]采用经典卷积神经网络 LeNet-5, 通过摄食强度自动分级方法来评估罗非鱼食欲。张佳林等^[1]采用 VAE-CNN 网络对鱼群摄食行为视频进行分类。VGGNet^[19]卷积神经网络的出现证明了增加网络深度可以提升模型效果, 但是加深网络深度会使模型体积变大且更复杂, 网络训练和预测需要更大的计算资源, 对硬件设备的要求更高。

相较于传统机器学习算法, 基于深度学习的鱼类摄食行为识别模型具有更高的识别精度和更好的鲁棒性。目前已有的大量研究均为实验室工况, 室外集约化养殖环境下基于移动设备和嵌入式设备的鱼类摄食行为研究较少。针对此种现象, 为实现室外圈养鲈鱼的精准投喂, 本研究提出了基于 MobileNetV3-Small 轻量级神经网络的鲈鱼摄食状态分类模型。

1 材料与方法

本研究于华中农业大学水产学院实践教学基地 (二) “零排放圈养系统” 搭建了圈养鲈鱼视频采集平台, 将一台海康威视摄像机 (DS-2CD3T47DWD-L) 安装固定在距离圈养塘边缘垂直高约 2.5 m 的位置进行拍摄 (图 1)。通过调整摄像机角度使其能够拍摄到圈养塘全部区域, 为避免过强的阳光反射影响拍摄效果, 摄像机镜头选择尽量背对太阳光源的视线方向。视频图像数据通过摄像机经 4G 路由器传输到本地录像机和广域网。所用摄像机、路由器和视频录像机均已做防水防晒处理。摄像机采集的原始视频图像分辨率为 2 560×1 440 像素, 帧率为 25 帧/s。图像处理语言为 Python3, 神经网络模型的搭建通过 Tensorflow2 实现。



1. 目标区域 2. 圈养塘 3. 摄像头 4. 录像机 5. 路由器
1. Target area 2. Captive pond 3. Camera 4. Video recorder 5. Wireless router

图 1 视频采集平台示意图

Fig.1 Video capture platform diagram

本研究中大口黑鲈 (*Micropterus salmoides*) 品种为“优鲈 1 号”, 已在该室外集约化环境下养殖超过 10 个月, 可以认为已适应该环境下的饲料投喂。试验用投喂饲料为浮性鲈鱼膨化配合饲料, 采购于荆门市五龙饲料有限公司, 饲料规格为 6#, 直径约 6 mm, 饲料密度小于

水的密度。随机取 300 粒饲料, 分为 3 组, 称质量后取均值得饲料百粒质量 23.2 g。圈养鲈鱼采用单场多轮投喂, 即每天投喂两场, 每场投喂若干轮, 投喂轮数由鱼群进食情况决定, 每轮投喂量为 500 g, 投喂时间点为上午 8:00 和下午 17:30。视频采集期间由水产养殖人员负责饱食投喂。视频图像开始采集时鲈鱼平均质量为 238.57 g, 养殖密度约为 20.04 kg/m³。为使本研究更好地适应室外真实养殖环境, 在视频采集过程中, 除暴雨、大风等恶劣天气外, 其他天气情况的视频均作为研究对象。

2 圈养鲈鱼数据集制作

通过投喂预试验发现, 在每轮饲料投喂量为 500 g 的情况下, 经观察所投喂饲料均漂浮于水面, 不会产生沉降且不破碎, 在每轮进食初始阶段鲈鱼抢食较激烈, 水面水花过多导致难以观察饲料剩余情况, 但在进食最后阶段抢食不激烈, 水面水花较少, 能够通过饲料剩余情况更清楚地观察当前鲈鱼进食情况。因此, 本研究随机统计的 7 d 内鲈鱼每轮进食时间如表 1 所示。在每轮投喂中, 将投喂动作结束至鲈鱼进食完成所用时间记为每轮进食时间。在每轮投喂中, 若鲈鱼将饲料全部摄食, 水面不存在或存在极少量 (小于 10 g) 剩余饲料, 则认为该轮进食已完成; 此外, 若水面存在部分剩余饲料且连续 5 s 内进食鲈鱼数量小于 5 也认为该轮鲈鱼进食完成。在同一场投喂过程中, 每轮进食时间逐渐增加, 说明鲈鱼食欲呈现下降趋势。通过计算得鲈鱼平均进食时间为 94 s, 此时水面基本处于稳定状态, 相对于其他时段能够更加清楚地观察水面饲料剩余情况, 分辨出鲈鱼的摄食状态, 为下一轮投喂进行决策。因此选择每轮投喂结束后第 80~110 秒的视频图像作为数据样本。

本研究参考文献[20-22]的鱼群摄食强度标准和鲈鱼在此试验条件下水面进食和饲料情况, 对鲈鱼在第 90 秒左右的食欲程度进行划分, 划分为“强”、“中”、“弱”、“无”4 个等级。其中“强”即在该时间段内水面已看不到鲈鱼和饲料, 表明该轮鲈鱼进食已经结束, 鲈鱼进食快, 食欲强; “中”即通过观察水面剩余饲料数量, 结合饲料百粒质量估算出对应干燥状态下剩余饲料的质量小于 100 g, 可视为鲈鱼食欲有所下降; “弱”即水面所剩饲料质量约大于 100 g, 可视为鲈鱼进食欲望不强烈; “无”即水面仍存在饲料, 但是没有鲈鱼进食, 表明鲈鱼已经没有食欲。鲈鱼食欲等级划分样本如图 2 所示。在实际投喂过程中, 由于塘内鲈鱼数量较多, 其中个体较大、性情凶猛的鲈鱼先进食, 个体较小且胆小的鲈鱼稍后进食。所以, 存在投喂 500g 饲料后, 鲈鱼进食部分饲料后整体食欲等级由“中” (“饥饿”状态) 变为“弱” (“非饥饿”状态), 但此时仍有个体较小和胆小的鲈鱼仍在进食, 剩余饲料仍会被缓慢地摄食完毕。为了使食欲等级的划分更有实际意义, 根据食欲等级将“强”和“中”图片划分为“饥饿”, “弱”和“无”划分为“非饥饿”数据集类别, 基于摄食状态的二分类结果更适用于实际养殖决策。

表 1 鲈鱼每轮进食时间和平均每轮进食时间
Table 1 Feeding time per round and average feeding time per round of perch

进食时间 Time of feeding	试验记录天数 Number of days of test records						
	第 1 天 Day 1	第 2 天 Day 2	第 3 天 Day 3	第 4 天 Day 4	第 5 天 Day 5	第 6 天 Day 6	第 7 天 Day 7
每轮进食时间 Feeding time per round	44	48	41	65	53	49	51
	40	47	75	68	60	67	59
	75	79	120	77	97	93	75
	97	85	102	103	90	112	101
	86	87	114	95	130	127	135
	115	115	112	135	42	45	56
	133	160	65	61	66	67	71
	60	64	86	63	82	86	86
	70	79	82	83	105	115	106
	93	88	95	105	117	140	128
	102	113	98	90	160	163	146
	126	122	128	122	-	-	-
平均每轮进食时间 Average feeding time per round	91	94	95	96	92	97	92

注：“-”表示该轮没有投喂。
Note:“-” means that no feed was fed for that round.

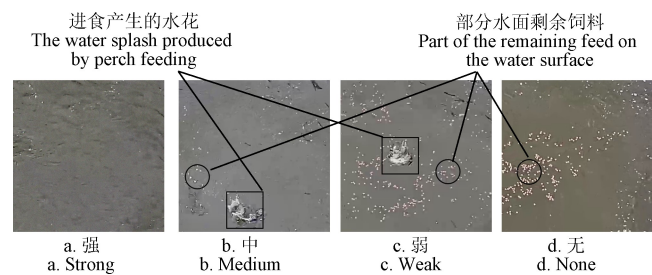


图 2 鲈鱼食欲等级
Fig.2 Perch appetite rating

将每轮投喂后第 80~110 秒的视频每隔 1 s 提取一帧图像，通过观察鱼群进食视频图像，选取能够准确反映鲈鱼进食和饵料状况的主要区域并进行裁剪，裁剪后图片分辨率为 512×512 像素，图片格式为.png。从中随机选取较具代表性的 10 000 张图片作为原始图片数据集，其中“饥饿”和“非饥饿”图片各有 5 000 张。原始图片数据集中训练集、验证集、测试集按照 6:2:2 比例随机划分。为了提高模型泛化能力，训练集采用离线随机数据增强的方式进行扩充，数据增强方式包括图片随机翻转、随机裁剪、添加高斯噪声。高斯噪声以均值为 0，方差为 0.001 进行添加。此外，以晴天、阴天和小雨为例，对室外不同天气情况进行了分析，发现鱼塘视频图像主要差异是光照不同。为减小不同天气对模型分类效果的影响，本研究选用图片色彩抖动方法对训练集数据进一步增强，使用的色彩抖动包括随机调整图片的饱和度、亮度和对比度。色彩抖动效果如图 3 所示，图片经色彩抖动增加了图片在不同天气中的情况，进一步提高模型的泛化能力。训练集经过数据增强后共 12 000 张图片，其中

“饥饿”和“非饥饿”图片各有 6 000 张，制作后的数据集如表 2 所示。

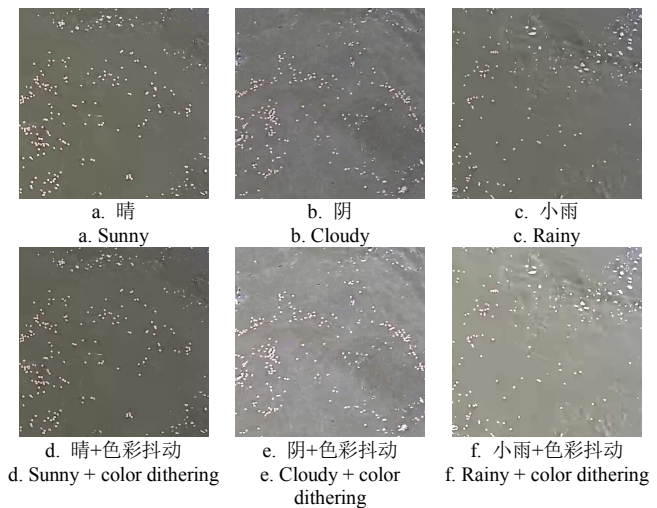


图 3 不同天气图片色彩抖动
Fig.3 Different weather picture color dithering

表 2 鲈鱼摄食状态数据集的制作
Table 2 Production of starvation status dataset for perch

数据集 Data set	数据集类别 Dataset category	样本数量 Number of sample	样本总数量 Total number of sample
训练集 Training set	饥饿	6 000	12 000
	非饥饿	6 000	
验证集 Validation set	饥饿	1 000	2 000
	非饥饿	1 000	
测试集 Test set	饥饿	1 000	2 000
	非饥饿	1 000	

3 基于机器视觉的鲈鱼摄食状态分类

3.1 基于 MobileNetV3-Small 的鲈鱼摄食状态分类

本研究选取 MobileNetV3-Small 作为鲈鱼摄食状态分类模型。MobileNetV3-Small 是 MobileNetV3^[21]系列中的一款轻量级神经网络,适用于低性能移动设备和嵌入式设备。本研究使用 512×512×3 图片作为 MobileNetV3-Small 模型输入。首先对输入图片进行标准卷积,卷积核大小为 3×3,卷积核个数为 16,激活函数为 h-swish,步长为 2。进一步将得到的 256×256×16 特征输入到一系列 block 单元中,其中包括 3 个卷积核为 3×3 深度卷积的 block 单元和 8 个卷积核为 5×5 深度卷积的 block 单元,得到大小为 16×16×96 的特征。然后通过 1×1 卷积核对输出的特征进行升维,得到 16×16×576 的特征。再通过全局平均池化^[22]使特征变为一维向量特征。最后经 2 个 1×1 的标准卷积后得到 2 个类别的概率值。MobileNetV3-Small 整体结构如表 3 所示。

表 3 MobileNetV3-Small 整体结构
Table 3 MobileNetV3-Small overall structure

输入 Input	操作 Operator	升维维度 Exp size	输出通道 #Out	注意力模 块 SE	非线性激 活函数 NL	步长 Stride
512 ² ×3	Conv2d, 3×3	-	16	-	HS	2
256 ² ×16	Bneck, 3×3	16	16	√	RE	2
128 ² ×16	Bneck, 3×3	72	24	-	RE	2
64 ² ×24	Bneck, 3×3	88	24	-	RE	1
64 ² ×24	Bneck, 5×5	96	40	√	HS	2
32 ² ×40	Bneck, 5×5	240	40	√	HS	1
32 ² ×40	Bneck, 5×5	240	40	√	HS	1
32 ² ×40	Bneck, 5×5	120	48	√	HS	1
32 ² ×48	Bneck, 5×5	144	48	√	HS	1
32 ² ×48	Bneck, 5×5	288	96	√	HS	2
16 ² ×96	Bneck, 5×5	576	96	√	HS	1
16 ² ×96	Bneck, 5×5	576	96	√	HS	1
16 ² ×96	Conv2d, 1×1	-	576	√	HS	1
16 ² ×576	Pool, 16×16	-	-	-	-	1
1 ² ×576	Conv2d, 1×1, NBN	-	1 024	-	HS	1
1 ² ×1 024	Conv2d, 1×1, NBN	-	2	-	-	1

注:“-”表示该层不能进行对应操作;“√”表示该层使用了注意力模块(SE);在操作(Operator)中,“Bneck”为 MobileNetV3 的 block 单元,“3×3”表示该层使用的卷积核大小,“NBN”表示该层卷积不使用批量归一化(Batch Normalization);升维维度(Exp size)表示每一个 Bneck 块中第一层 1×1 conv2d 的卷积核的个数;激活函数(NL)表示每层所使用的非线性激活函数,其中“HS”表示 h-swish 函数,“RE”表示使用 ReLU 激活函数。

Notes:“-” denotes that the layer cannot perform the corresponding operation;“√” denotes that the layer uses the attention module (SE); in the operation (Operator), “Bneck” is the block unit of MobileNetV3, “3×3” denotes the size of the convolutional kernel used in the layer, “NBN” denotes that the layer does not use Batch Normalization for convolution; the ascending dimension (Exp size) denotes the number of convolution kernels in the first layer of 1×1 conv2d in each Bneck block; nonlinearity (NL) denotes the nonlinear activation function used in each layer, where “HS” is h-swish function and “RE” is ReLU activation function.

MobileNetV3-Small 在基本 block 单元的设计上结合了 MobileNetV1^[23]中的深度可分离卷积、MobileNetV2^[24]线性瓶颈的逆残差结构和 Squeeze and Excite^[25] (SE) 轻量级的注意力结构。在 MobileNetV3-Small 的 block 单元中,先通过一个 1×1 卷积层对输入特征进行升维,将升

维后的特征进行卷积核为 3×3 的深度卷积,然后将输出特征输入到 SE 模块。在 SE 模块中,首先对特征的每个通道进行全局平均池化,得到 1×1 的 exp size 向量,再将该向量通入第一层全连接层(Fully Connected layers, FC),激活函数为 ReLU,得到输出大小为 1×1×(exp size/4)的张量。然后将输出通入第二层 FC,激活函数为 h-swish,得到 1×1×exp size 的向量,将该向量中的每个元素乘以输入到 SE 层特征的对应通道特征矩阵,得到经 SE 模块加权之后的特征。最后采用 1×1 卷积对特征降维,输出维度如表 4 输出通道所示。MobileNetV3-Small 基本 block 单元如图 4 所示。

表 4 各模型性能指标对比结果

Table 4 Comparison results of performance indicators of each model

模型 Model	准确率 Accuracy/ %	召回率 Recall/ %	精准率 Precision/ %	F1 分数 F1 score/%	FLOPs/ M	平均速率 Average rate/(帧·s ⁻¹)
KNN (K-Nearest Neighbors)	86.86	88.02	96.19	87.10	-	37.65
SVM (Support Vector Machine)	75.50	71.58	77.99	74.65	-	38.03
GBDT (Gradient Boosting Decision Tree)	96.85	97.59	96.23	96.90	-	37.73
Stacking	96.82	97.23	96.48	96.86	-	38.67
ResNet-18	99.60	99.90	99.30	99.60	18 959	27.36
ShuffleNetV2	98.95	99.39	98.50	98.94	1 507	29.67
MobileNetV3-Large	99.20	98.81	99.60	99.20	2 144	25.77
MobileNetV3-Small	99.60	99.40	99.80	99.60	582	39.21

注:“-”表示未对模型进行相应计算。FLOPs 为计算量。

Note:“-” indicates that the corresponding calculation was not performed for the model. FLOPs is floating point operations.

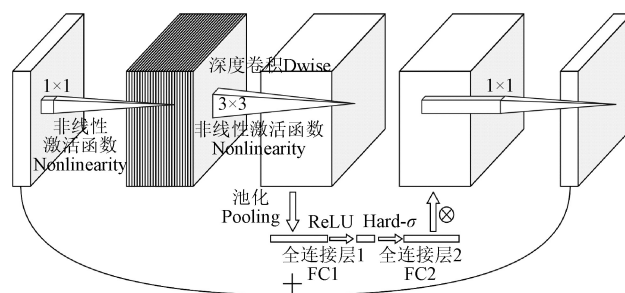


图 4 MobileNetV3-Small 基本 block 单元

Fig.4 MobileNetV3-Small basic block unit

MobileNetV3-Small 在制作的鲈鱼摄食状态分类数据集上进行训练。训练选取批大小为 16,学习率为 0.000 1,迭代次数为 50,最终得到训练集和验证集的损失值和准确率曲线如图 5 所示。

由图 5 可见迭代次数达到 10 次后,训练集损失值曲线逐渐平缓,说明模型已经接近收敛,最终在验证集上的准确率达到 99.7%。训练好的最终模型在测试集上的准确率达到 99.60%,召回率为 99.40%,精准率为 99.80%,F1 分数为 99.60%。基于 MobileNetV3-Small 模型能够有效地对圈养鲈鱼摄食状态进行分类,可以为鲈鱼的高效投喂提供理论依据。

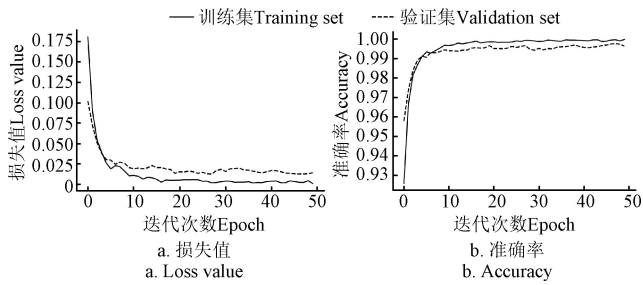


图 5 MobileNetV3-Small 损失值和准确率曲线

Fig.5 MobileNetV3-Small loss value and accuracy curve

3.2 MobileNetV3-Small 与其他分类模型对比

为验证 MobileNetV3-Small 模型的分类效果,本研究选用传统机器学习模型与其进行对比。传统机器学习选用了 KNN^[26]、SVM^[27]和 GBDT^[28]模型和基于 3 种模型的 Stacking^[29]集成学习模型。传统机器学习中提取的图片特征包括纹理、颜色和形状共 28 组特征。纹理特征包括图像的灰度图像特征、不变矩特征和灰度共生矩阵特征共 17 组特征。颜色特征是将 RGB 图片转到 HSV 空间后,在 HSV 空间上的一阶、二阶和三阶颜色矩的均值、方差和偏移量共 9 组特征。形状特征是通过 Canny 算子得到图像轮廓,然后计算轮廓面积和周长得到 2 组特征。为保证所提取特征的有效性,通过特征工程对特征进行筛选。首先采用皮尔森相关系数计算出各组特征与分类结果的相关性,得到对分类结果影响较大的 13 组特征。进一步通过皮尔逊相关系数计算出各组特征之间的相关性矩阵,剔除特征之间相关性较高的特征,最终得到 6 组特征。利用得到的 6 组特征对 KNN、SVM、GBDT 和 Stacking 集成学习模型进行训练,其中 KNN 和 SVM 模型采用网格搜索的方式寻找最优参数。此外,为进一步验证选取模型的合理性,通过 MobileNetV3-Small 与 ResNet-18、轻量级网络 ShuffleNetV2 和 MobileNetV3-Large 深度学习模型进行对比。

本研究通过准确率 (Accuracy)、召回率 (Recall)、精准率 (Precision)、F1 分数 (F1)、模型计算量 (Floating Point Operations, FLOPs) 和平均分类速率 6 个指标对各个模型进行评估,其计算公式和定义如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (3)$$

$$\text{F1} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \times 100\% \quad (4)$$

式中 TP 为真阳性 (True Positive), 在本文中模型输出类别为饥饿, 数据集标签也为饥饿; FP 为假阳性 (False Positive), 即模型输出类别为饥饿, 数据集标注为非饥饿; TN 为真阴性 (True Negative), 即模型输出类别为非饥饿, 数据集标注也为非饥饿; FN 为假阴性 (False Negative), 即模型输出类别为饥饿, 数据集标注为非饥饿。模型计算量 (Floating Point Operations, FLOPs) 是模型浮点运算数,

可以用来衡量模型的复杂度。平均分类速率是将 1 000 张图片进行分类, 得到的平均每秒分类图片的数量。

为保证结果公平性, 几种模型均使用同一数据集。深度学习模型选取批大小为 16, 学习率为 0.000 1, 迭代次数为 50, 优化器为 Adam 优化器。试验所用计算机运行环境为 Win10 Inter(R)Core(TM)i7-10750H CPU @2.60 GHz, 运行内存 32 G。各个深度学习模型训练在经过 50 个 epochs 后均已收敛, 得到 MobileNetV3-Small 与其他模型对比结果如表 4 所示。

基于传统机器学习模型的鲈鱼摄食状态分类在平均分类速率上与 MobileNetV3-Small 神经网络模型相差不大。但 MobileNetV3-Small 神经网络模型的准确率分别高出 KNN、SVM、GBDT 和 Stacking 模型 12.74、23.85、3.60 和 2.78 个百分点。可能因为室外环境视频图片噪声干扰较多, 人工提取和筛选特征比较困难。而基于神经网络模型的自动特征提取能够更有效地提取有用的特征, 分类效果更好。

通过对比 ResNet-18、ShuffleNetV2、MobileNetV3-Large 和 MobileNetV3-Small, 结果表明, 4 种深度学习模型的准确率、召回率、精准率、F1 分数指标较为相近, 且高于传统机器学习模型。上述 4 种深度学习模型均能实现鲈鱼饥饿状态分类。但在水产养殖中, 嵌入式硬件设备计算能力的限制要求模型有更小的计算量和更快的分类速率。MobileNetV3-Small 模型的计算量为 582 M, 平均分类速率为 39.21 帧/s。在其他 3 种模型中, 计算量最优的 ShuffleNetV2 仍是 MobileNetV3-Small 模型的 2.6 倍; 平均分类速率最优的 ShuffleNetV2 模型约为 MobileNetV3-Small 模型平均分类速率的 3/4。结合模型计算量和平均分类速率, MobileNetV3-Small 分类效果优于其他模型。综上, 本研究选取 MobileNetV3-Small 作为鲈鱼摄食状态分类模型。

4 圈养鲈鱼投喂对比试验

为验证基于 MobileNetV3-Small 的鲈鱼摄食状态分类模型的实际效果, 本研究设计了圈养鲈鱼投喂对比试验, 试验在华中农业大学水产学院教学实习基地 (二) “零排放圈养模式” 池塘进行, 如图 6 所示, 3 号圈养塘作为试验组, 6 号圈养塘作为对照组, 2 个圈养塘处于同一养殖环境下。



1. 6 号圈养塘 (对照组) 2. 3 号圈养塘 (试验组) 3. 摄像头 4. 录像机与路由器
1. No.6 fish pond (Control group) 2. No.3 fish pond (Test group) 3. Camera 4. Video recorder and wireless router

图 6 鲈鱼投喂对比试验场地

Fig.6 Perch feeding comparison test site

试验组鱼塘采用 MobileNetV3-Small 神经网络模型输出结果进行决策投喂, 对照组则通过水产养殖人员根据养殖经验进行投喂。试验采用单场多轮投喂方式, 每轮饲料投喂量均为 500 g, 2 个圈养塘除投喂决策方式不同外, 其他条件保持一致。试验组视频获取方式为 Jupyter Nootbook 通过本地直连摄像机的方式获取实时视频图像, 视频图像原始分辨率为 2 560×1 440 像素, 截取同训练集指定区域后图像分辨率为 512×512 像素。试验组投喂时, 选取每轮投喂后第 90 秒、第 92 秒、第 94 秒、第 96 秒和第 98 秒的水面图片, 将 5 帧图片作为鲈鱼投喂决策模型的输入, 输出结果为每张图片对应的摄食状态, 并在该轮投喂结束后第 100 秒进行投喂决策。在 5 帧图片输出结果中, 若预测为“饥饿”的次数占比不小于总数的 60%, 则直接进行下一轮投喂; 否则此轮不投喂, 经 50 s 后重新采集图片进行投喂决策, 若决策结果仍为不投喂, 则结束本场投喂。基于多张图片决策减少了单张图片分类错误导致决策错误的情况。试验组投喂决策根据模型输出结果进行投喂, 投喂决策路线图如图 7 所示。

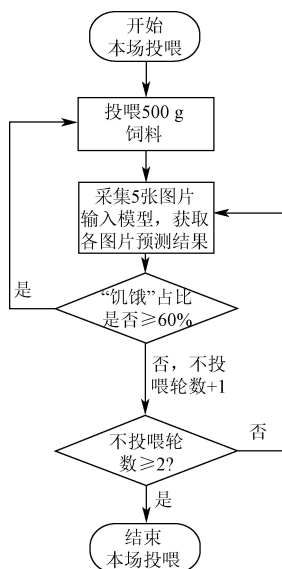


图 7 投喂决策路线图
Fig.7 Feeding decision roadmap

对比投喂试验选用饵料系数 (Feed Conversion Ratio, FCR) 和质量增加率 (Weight Gain Ratio, WGR) 作为衡量两组圈养塘鲈鱼生长情况的指标。饵料系数能反映增加单位水产品产量所用的饲料量, 饵料系数比值越小, 表示增加单位水产品产量所用的饲料越少。质量增加率表示一段时间内水产品增加质量与初始质量的关系, 能够直观地反应该段时间内鲈鱼的质量增加情况。饵料系数和质量增加率计算公式如下:

$$FCR = \frac{L_s}{L_1 - L_0} \quad (5)$$

$$WGR = \frac{L_1 - L_0}{L_0} \times 100\% \quad (6)$$

式中 L_s 为该段时间内投喂的饲料总量, kg; L_0 为投喂开始时鲈鱼的总质量, kg; L_1 为投喂结束时鲈鱼的总质量, kg。本试验开始到养殖结束期间, 鲈鱼已适应养殖环境,

损失数量可以忽略不计, 因此使用养殖结束时统计的鲈鱼数量作为试验期间鲈鱼数量。鲈鱼平均质量采用简单随机抽样的方式进行统计, 每次抽样数量为 30 尾, 计算平均质量。投喂 14 d 后, 对比 2 组试验结果如表 5 所示。

表 5 鲈鱼投喂对比试验结果
Table 5 Results of comparison test of perch feeding

组别 Group	数量 Number/ 尾	初始质量 Initial weight L_0 /kg	最终质量 Final weight L_1 /kg	投喂量 Feed weight L_s /kg	饵料系数 Feed Conversion Ratio (FCR)	质量增加率 Weight Gain Ratio (WGR)/%
试验组 Test group	2 096	660.24	696.92	52.00	1.42	5.56
对照组 Control group	2 106	694.98	714.98	54.50	2.72	2.88

试验期间试验组圈养塘水质参数 pH 值变化范围在 7.12~7.90, 氨氮变化范围在 0.54~0.72 mg/L, 参考国家渔业水质标准^[30], 2 个参数均处于合理范围, 说明试验组圈养塘水质未发生明显下降现象。由对比试验结果得出, 试验组圈养塘饵料系数为 1.44, 对照组饵料系数为 2.72, 试验组饵料系数为对照组的 52.20%, 该段时间内试验组饵料利用率高于对照组。试验组质量增加率为 5.56%, 对照组质量增加率为 2.88%, 试验组质量增加率提高 2.68 个百分点, 该段时间内试验组鲈鱼生长速率高于对照组。综上, 试验组圈养塘的饵料系数和质量增加率均优于对照组, 说明本试验投喂模型能够减少饲料浪费, 节约养殖成本, 提高养殖效率。基于 MobileNetV3-Small 模型的鲈鱼摄食状态分类有较好的应用效果, 基于该模型的投喂决策在一定程度上能够代替养殖人员投喂决策, 为室外集约化水产养殖提供一种有效参考。

5 结 论

针对室外养殖环境复杂、鲈鱼饥饿状态难以辨别等特点, 本研究提出了基于 MobileNetV3-Small 的圈养鲈鱼摄食状态分类模型, 在真实室外养殖环境下达到了较好的分类投喂效果, 为室外集约化养殖环境下的智能投喂提供了参考。

1) 依据鲈鱼食欲等级划分了鲈鱼摄食状态数据集。MobileNetV3-Small 网络模型在该数据集上的准确率达到 99.60%, 召回率为 99.40%, 精准率为 99.80%, F1 分数为 99.60%。

2) 通过与 ResNet-18, ShuffleNetV2 和 MobileNetV3-Large 深度学习模型相比, MobileNetV3-Small 模型的计算量最小为 582 M, 平均分类速率最大为 39.21 帧/s。与 KNN、SVM、GBDT 和 Stacking 传统机器学习模型相比, MobileNetV3-Small 模型的综合准确率高出 12.74、23.85、3.60 和 2.78 个百分点。

3) 通过圈养鲈鱼投喂试验表明, 基于该分类模型决策的投喂方式为人工投喂方式饵料系数的 52.20%, 质量增加率提高 2.68 个百分点。

[参 考 文 献]

[1] 张佳林, 徐立鸿, 刘世晶. 基于水下机器视觉的大西洋鲑

- 摄食行为分类[J]. 农业工程学报, 2020, 36(13): 158-164.
- Zhang Jialin, Xu Lihong, Liu Shijing. Classification of Atlantic salmon feeding behavior based on underwater machine vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(13): 158-164. (in Chinese with English abstract)
- [2] 郭强. 基于计算机视觉的循环水养殖镜鲤的摄食状态检测方法研究[D]. 上海: 上海海洋大学, 2018.
- Guo Qiang. Fish Feeding Behavior Detection Method Under Circulating Aquaculture Based on Computer Vision[D]. Shanghai: ShangHai Ocean University, 2018. (in Chinese with English abstract)
- [3] Barraza-Guardado R H, Martínez-Córdova L R, Enríquez-Ocaña L F, et al. Effect of shrimp farm effluent on water and sediment quality parameters off the coast of Sonora, Mexico[J]. Ciencias Marinas, 2014, 40(4): 221-235.
- [4] Papadakis V M, Papadakis I E, Lamprianidou F, et al. A computer-vision system and methodology for the analysis of fish behavior[J]. Aquacultural Engineering, 2012, 46: 53-59.
- [5] Pautsina A, Císař P, Štys D, et al. Infrared reflection system for indoor 3D tracking of fish[J]. Aquacultural Engineering, 2015, 69: 7-17.
- [6] 乔峰, 郑堤, 胡利永, 等. 基于机器视觉实时决策的智能投饵系统研究[J]. 工程设计学报, 2015, 22(6): 528-533.
- Qiao Feng, Zheng Ti, Hu Liyong, et al. Research on smart bait casting machine based on machine vision technology[J]. Chinese Journal of Engineering Design, 2015, 22(6): 528-533. (in Chinese with English abstract)
- [7] 郭俊. 基于图像与声音信息的养殖鱼群摄食规律与投饵技术研究[D]. 宁波: 宁波大学, 2018.
- Guo Jun. Research on Feeding Patterns and Bait Technology of Fish Culture Based on Information of Image and Sound[D]. Ningbo: Ningbo University, 2018. (in Chinese with English abstract)
- [8] Atoum Y, Srivastava S, Liu X M. Automatic feeding control for dense aquaculture fish tanks[J]. IEEE Signal Processing Letters, 2014, 22(8): 1089-1093.
- [9] 陈明, 张重阳, 冯国富, 等. 基于特征加权融合鱼类摄食活动强度评估方法[J]. 农业机械学报, 2020, 51(2): 245-253.
- Chen Ming, Zhang Chongyang, Feng Guofu, et al. Intensity assessment method of fish feeding activities based on feature weighted fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 245-253. (in Chinese with English abstract)
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [11] Chen L, Yang X T, Sun C H, et al. Feed intake prediction model for group fish using the MEA-BP neural network in intensive aquaculture[J]. Information Processing in Agriculture, 2020, 7(2): 261-271.
- [12] Adegboye M A, Aibinu A M, Kolo J G, et al. Incorporating intelligence in fish feeding system for dispensing feed based on fish feeding intensity[J]. IEEE Access, 2020, 8: 91948-91960.
- [13] Måløy H, Aamodt A, Misimi E. A spatio-temporal recurrent network for salmon feeding action recognition from underwater videos in aquaculture[J]. Computers and Electronics in Agriculture, 2019, 167: 105087.
- [14] Villon S, Mouillot D, Chaumont M, et al. A deep learning method for accurate and fast identification of coral reef fishes in underwater images[J]. Ecological Informatics, 2018, 48: 238-244.
- [15] 赵德安, 刘晓洋, 孙月平, 等. 基于机器视觉的水下河蟹识别方法[J]. 农业机械学报, 2019, 50(3): 151-158.
- Zhao Dean, Liu Xiaoyang, Sun Yueping, et al. Detection of underwater crabs based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 151-158. (in Chinese with English abstract)
- [16] Zhang S, Yang X T, Wang Yi Z, et al. Automatic fish population counting by machine vision and a hybrid deep neural network model[J]. Animals, 2020, 10(2): 364.
- [17] Tseng C H, Hsieh C L, Kuo Y F. Automatic measurement of the body length of harvested fish using convolutional neural networks[J]. Biosystems Engineering, 2020, 189: 36-47.
- [18] Zhou C, Xu D M, Chen L, et al. Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision[J]. Aquaculture, 2019, 507: 457-465.
- [19] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the IEEE International Conference on Learning Representations (ICLR). San Diego: IEEE, 2015.
- [20] Øverli Ø, Sørensen C, Nilsson G E. Behavioral indicators of stress-coping style in rainbow trout: do males and females react differently to novelty?[J]. Physiology & Behavior, 2006, 87(3): 506-512.
- [21] Howard A, Sandler M, Chu G, et al. Searching for mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019.
- [22] Lin M, Chen Q, Yan S. Network in network[C]//Proceedings of the IEEE International Conference on Learning Representations (ICLR). Banff: IEEE, 2014.
- [23] Howard A G, Zhu M, Chen B, et al. MobileNets efficient convolutional neural networks for mobile vision applications[C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Honolulu: IEEE, 2017.
- [24] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Salt Lake: IEEE, 2018.
- [25] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Salt Lake: IEEE, 2018.
- [26] Abeywickrama T, Cheema M A, Taniar D. K-nearest neighbors on road networks: A journey in experimentation and in-memory implementation[J]. Proceedings of the VLDB Endowment, 2016, 9(6): 492-503.
- [27] Hui Y, Mei X S, Jiang G D, et al. Milling tool wear state recognition by vibration signal using a stacked generalization ensemble model[J]. Shock and Vibration, 2019, 2019(3): 1-16.
- [28] Friedman J H. Greedy function approximation: A gradient boosting machine[J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [29] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 183-185.
- [30] 国家环境保护局. 渔业水质标准: GB 11607-1989[S]. 北京: 中国标准出版社, 1990.

Classification of perch ingesting condition using lightweight neural network MobileNetV3-Small

Zhu Ming^{1,3}, Zhang Zhenfu^{1,2}, Huang Huang^{1,2*}, Chen Yanyan¹, Liu Yadong¹, Dong Tao¹

(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; 3. Engineering Research Center of Green Development for Conventional Aquatic Biological Industry in the Yangtze River Economic Belt, Ministry of Education, Wuhan 430070, China)

Abstract: Intelligent feeding has widely been used to determine the amount of feed from a smart prediction about the hunger degree of fish, thereby effectively reducing the waste of feed in the modern aquaculture industry, especially for outdoor intensive fish breeding environments. However, redundant data collected by mobile monitoring devices has caused a huge calculation load for most control systems. An accurate classification of the hunger degree of fish still remains an unsolved problem. Taking the captive perch as the tested object, this work aims to design an image capture system for the perch feeding using MobileNetV3-Small of lightweight neural network. The system also consisted of 2 captive ponds, a camera, and a video recorder. In the test, 4202 perches were randomly fed with adequate or inadequate feed, where a camera was selected to record the water surface every day. 10 000 images were collected after 2-week monitoring to record the perch ingesting condition in the period of 80~110 seconds after per round feeding condition, where 50% belonged to “hungry” condition, and the rest was “non-hungry” condition. These initial images were then divided as training, validation, and testing set, according to a rate of 6:2:2. Four image processing operations were applied on the training set, containing random flipping, random cropping, adding Gaussian noise, and color dithering, thereby expanding the training set from 6 000 to 12 000 images. As such, the more generalized model greatly enhanced the image features and training samples. Next, a MobileNetV3-Small of lightweight Neural Network was selected to classify the ingesting condition of perches. The model was trained, tested, and established on the Tensorflow2 platform, where the images of the training set were selected as the input, whereas, the ingesting condition as the output. Finally, a 2-week feeding contrast test was carried out in the outdoor culture environment to verify the accuracy of the model. Two groups were set for 4202 perches in this test, 2096 of the test group and 2106 of the control group, where the amount of feed was determined according to the classification of model and conventional experience. Meanwhile, the total mass and quantity of the two groups were recorded at the beginning and end of the test, as well as the total amount of consumed feed. Correspondingly, it was found that the MobileNetV3-Small network model achieved a combined accuracy of 99.60% in the test set with an F1 score of 99.60%. The MobileNetV3-Small model presented the smallest Floating Point Operations of 582 M and the largest average classification rate of 39.21 frames/s, compared with ResNet-18, ShuffleNetV2, and MobileNetV3-Large deep learning models. Specifically, the combined accuracies of the MobileNetV3-Small model were 12.74, 23.85, 3.6, and 2.78 percentage points higher than that of the traditional machine learning models KNN, SVM, GBDT, and Stacking. Furthermore, the test group of perch was achieved a lower Feed Conversion Ratio of 1.42, and a higher Weight Gain Ratio of 5.56%, compared with the control group, indicating that the MobileNetV3-Small model performed a better classification on the ingesting condition in a real outdoor culture environment. Consequently, the classification of the ingesting condition can widely be expected for the efficient decision-making for the amount of fish feed, particularly suitable for the growth of fish. The finding can provide a further reference for efficient and intelligent feeding in an intensive cultural environment.

Keywords: aquaculture; machine vision; image recognition; deep learning; neural network; perch