

基于 ResNet – CA 的鱼群饱腹程度识别方法

孙龙清^{1,2} 王新龙^{1,2} 王泊宁^{1,2} 王嘉煜^{1,2} 孟新宇^{1,2}

(1. 中国农业大学信息与电气工程学院, 北京 100083;
2. 农业农村部国家数字渔业创新中心, 北京 100083)

摘要: 投喂作为水产养殖过程中的一个关键环节, 饵料的投喂量直接影响水产品的质量和养殖成本。然而, 目前的投喂方法包括人工投喂和机器定时定量投喂, 大多依靠人工经验, 很难实现精准投喂。本文基于改进的 ResNet34 识别鱼群不同的饱腹程度。根据鱼群在不同饱腹阶段表现的摄食行为创建了含有 5 种不同饱腹程度的数据集, 并采用数据增强操作对图像进行预处理。其次在原始模型 ResNet34 的基础上, 本文提出使用坐标注意力机制, 使模型在对图像进行特征提取的过程中能够做到专注于更大区域范围。并且使用深度可分离卷积的方式来代替传统卷积, 减少模型参数量。为了评估改进的有效性, 分析了改进后的模型在鱼群饱腹程度数据集上的性能, 并将其与原模型 ResNet34、AlexNet、VGG16、MobileNet – v2、GoogLeNet 等经典卷积神经网络架构进行比较。综合实验结果表明, 该模型相较于原模型参数量减少 46.7%, 准确率达到 93.4%, 相较于原模型提升 3.4 个百分点, 同时改进后的模型在准确率、精确度、召回率等方面也都优于其他卷积神经网络。综上所述, 本模型实现了性能与参数量之间的良好平衡, 为后续模型在实际养殖环境中的部署并指导养殖户改善和制定投喂策略提供了可能。

关键词: 鱼群; 摄食行为; 注意力机制; 深度可分离卷积; 卷积神经网络; 饱腹程度

中图分类号: TP391 **文献标识码:** A **文章编号:** 1000-1298(2022)S2-0219-07

Identification Method of Fish Satiation Level Based on ResNet – CA

SUN Longqing^{1,2} WANG Xinlong^{1,2} WANG Boning^{1,2} WANG Jiayu^{1,2} MENG Xinyu^{1,2}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. National Digital Fisheries Innovation Center, Ministry of Agriculture and Rural Affairs, Beijing 100083, China)

Abstract: Feeding as a key part of the aquaculture process, the amount of bait fed directly affects the quality of aquatic products and the cost of aquaculture. However, the current feeding methods include manual feeding and machine feeding at regular intervals, which mostly rely on manual experience and are difficult to achieve accurate feeding. Different satiation levels of fish were identified based on the improved ResNet34, which was important for achieving accurate control of bait feeding in the future. A dataset containing five different satiation levels was created based on the feeding behaviors exhibited by fish at different satiation stages, and the images were pre-processed using data enhancement operations. Secondly, based on the original model ResNet34, the use of coordinate attention mechanism was proposed to enable the model to focus on a large area in the process of feature extraction of images. And the depth-separable convolution was used instead of the traditional convolution to reduce the number of model parameters. To evaluate the effectiveness of the improvements, the performance of the improved model was analyzed on the fish satiation dataset and compared it with the original model ResNet34, AlexNet, VGG16, MobileNet – v2, GoogLeNet and other classical convolutional neural network architectures. The comprehensive experimental results showed that the model reduced the amount of parameters by 46.7% and achieved an accuracy of 93.4% compared with the original model, which had a 3.4 percentage points improvement compared with the original model, and the improved model also outperformed other convolutional neural networks in terms of accuracy, precision, recall, and F1 score. In summary, the model achieved a good balance between performance and number of participants, which provided the possibility for subsequent models to be deployed in real farming environments and guide farmers in

improving and developing feeding strategies.

Key words: school of fish; feeding behavior; attention mechanism; deep separable convolution; convolutional neural network; satiation level

0 引言

中国虽是一个水产养殖大国,但发展模式却比较落后,自动化水平低^[1],饵料投喂不精确等。水产养殖过程中,投喂是至关重要的一个环节^[2],饲料约占养殖水产品总成本的 50% ~ 80%^[3-4],合理的投喂方式能够有效减少饵料浪费,提高养殖效益。目前,我国主要采用传统的人工投喂和机器定时投喂这两种模式进行投喂^[5],在实际的投喂过程中容易忽略环境变化等因素带来的影响^[6-7],导致投喂不足或过量^[8]。投喂不足会造成鱼类长时间处于饥饿状态,抵抗力下降,生长缓慢,产量减少等问题;若投喂过量则会导致饵料浪费、水质污染等问题^[9-11]。

在摄食过程中随着鱼群饱腹程度的增加,鱼群的摄食欲望也会降低,导致鱼群摄食行为发生变化^[12]。因此,分析鱼群摄食行为,确定鱼群的饱腹程度来指导精准投喂,既满足鱼类健康生长的需求,又不过量投饵,是当前水产养殖过程中急需解决的问题,同时也是迈向智能化养殖的必由之路。

目前,在摄食行为监测和分析领域主要以机器视觉和声学为主。在基于机器视觉的摄食行为分析中,贾成功等^[13]对采集的鱼群摄食图像进行处理,获取个体鱼坐标和鱼群面积特征,通过建模分析鱼群状态并搭建智能投饵系统实现合理投喂。YE 等^[14]通过提取速度等特征对鱼群的行为进行分析,提出了一种基于鱼类行为的摄食评估方法来合理控制投喂,实现福利化养殖。ZHAO 等^[15]通过提取鱼群行为特征,提出了一些新的评估方法来实现水产养殖过程中鱼群食欲的实时变化,指导投喂。上述通过机器学习的方法,一定程度上完成了对鱼群的智能投喂,但是仍然存在一定的问题,研究提到的坐标、面积、速度等指标都需要进行量化分析,真实水产养殖环境下养殖密度高,上述指标难以精准量化,为后续精准投喂造成困难^[16]。

在基于机器声学的摄食行为分析中荆丹翔等^[17]使用成像声呐对鱼群信息进行采集,通过对采集的数据进行图像处理等一系列操作后可以根据连续的两帧图像来锁定目标的空间位置,从而获得多个目标的运动轨迹,为鱼群的摄食行为监测提供了新的技术方案。MALLEKH 等^[18]利用声音传感器和接收器监测鱼类在摄食过程中发出的声

音来分析鱼类的摄食活动,有效地克服了在水质浑浊的情况下观察效果不好的问题。KOLAREVIC 等^[19]将声学加速传感器植入到鱼体内,可以实时监测鱼类的行为变化,通过监测指标的变化来量化鱼类的饥饱程度。而上述提到的使用声学方法来对鱼群摄食行为进行监测的研究,虽然克服了当鱼群密度过高时所带来的鱼体目标难以追踪,鱼类行为难以量化等问题,但是却需要传感器的支持,传感器需要安装在水环境中,难免会对鱼类生存产生影响,并且传感器价格昂贵、易损坏^[20],因此以声学为基础的方法难以在实际的养殖环境中大规模使用。

本文提出一种基于深度学习的非量化鱼类行为分析方法,通过神经网络自动完成数据特征与表达关系的自学习^[21],避免鱼类行为难以量化的问题,并且通过引入坐标注意力机制获取更大的特征提取范围,使用深度可分离卷积^[22]方法代替传统卷积来减小模型体积,从而改进 ResNet-34^[23],实现对 5 种不同鱼群饱腹程度的识别,并为后续鱼群饵料的精准投喂提供基础。

1 材料与方法

1.1 视频数据获取

实验数据采集地址为山东省烟台市某水产有限公司实验车间。车间内养殖池直径约 3.2 m,池内温度维持在 22 ~ 25℃^[24],另外养殖水体的溶解氧、pH 值、氨氮等水环境因素采用传感器实时监测并且都维持在最适范围内。由于斑石鲷生长速度快,适应性强,具有较高经济价值,因此选择斑石鲷作为研究对象。选取斑石鲷 50 尾,月龄 4 个月。在开展实验前,实验鱼已在养殖池暂养 7 d,使其充分适应实验环境,避免应激影响数据采集。视频采集系统主要是为了获取鱼群的视频数据,该系统主要由计算机、支架摄像头、光源、传感器等组成,如图 1 所示。摄像头采用海康威视摄像机,800 万像素,分辨率可达 3 840 像素×2 160 像素,摄像头通过支架安装在养殖池斜上方。因为本实验仅需要鱼群摄食过程中的数据,所以着重对摄食过程以及摄食前后 10 min 进行录制,为了保证视频流畅性,录制过程中以速度 25 f/s 进行视频数据采集。

1.2 图像数据预处理

图像数据的获取是从上述采集好的视频数据中

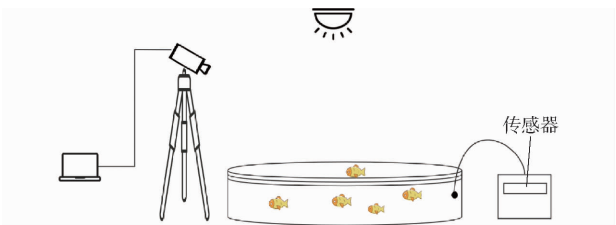


图 1 视频采集系统图

Fig. 1 Video capture system diagram

进行图像帧的提取,因为视频中相邻帧图像特征相似度极高,所以本文采用差帧提取的方式进行图像数据的获取,借助 OpenCV 来完成图像帧的自动提取,并通过 Python 脚本实现对所提取图像帧的差帧保存,帧间差为 50 帧,获得图像数据。获取图像数据后需要对数据进行筛选和标注,在获取的图像数据里面会有一些图像光照不均匀,存在外来噪声等问题,首先要将这些数据清理掉,避免这些无效数据影响后续实验。按照养殖专家的建议,本研究将鱼群饱腹程度划分为 5 个等级,如表 1 所示。

表 1 鱼群饱腹程度划分

Tab. 1 Classification of fish satiety level

饱腹程度	特征
饥饿	鱼群迅速向投饵区聚集,在饵料之间移动,并吃掉所有饵料
三成饱	游速变慢、聚集度降低
七成饱	少量鱼存在吐食现象,但是没有剩余饵料
饱食	部分斑石鲷摄食后会立刻将饵料吐出来,并存在少量剩余饵料
过饱	鱼对饵料没有反应,池底有剩余饵料

实验按照表 1 分类标准对图像数据集的标注进行严格划分,保证试验数据的高质量。整理后共获得分辨率 3 840 像素 × 2 160 像素图像 1 415 幅,其中饥饿 290 幅、三成饱 264 幅、七成饱 256 幅、饱食 276 幅、过饱 329 幅。并按照 7:3 划分为训练集和测试集。部分样本图像如图 2 所示。

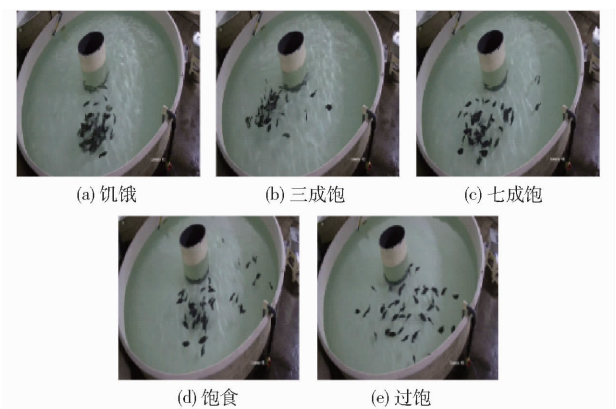


图 2 不同饱腹程度下的摄食图像

Fig. 2 Images of ingestion at different levels of satiety

2 模型结构

2.1 ResNet 模型

在 ResNet 网络提出以前,大多数的模型层数较少,有研究人员尝试堆叠网络层数,以达到更好的结果。但随着网络层数的堆叠,出现了两个问题:堆叠层数之后的网络难以收敛,因为梯度爆炸或者消失问题一开始就在阻碍网络的收敛;即使当更深的网络开始收敛时,却出现了退化问题,随着网络深度的增加,准确率达到饱和然后迅速下降。使得最终结果甚至不如浅层网络。HE 等^[23]提出的 ResNet 模型,其结构参数如表 2 所示,并且该模型在堆叠网络层数的基础上针对上述两种问题提供了解决方法。首先解决了梯度爆炸、消失问题;假设拟合目标函数为 $H(x)$,非线性叠加层为 $F(x)$,以前的网络模型是使 $F(x)$ 逼近 $H(x)$,但是随着网络层数的增加,较难很好地逼近 $H(x)$,因此模型使用残差块,结构如图 3 所示,通过快捷连接 (Shortcut connection) 的方式,使得 $F(x)$ 逼近 $H(x) - x$,让浅层特征恒等映射到深层,并且在反向传播的过程中深层梯度可以直接传回浅层,有效避免了网络退化问题的发生。最终通过不断堆叠网络层数,获得深度较大的网络模型和更好的网络性能。

表 2 ResNet34 网络参数

Tab. 2 ResNet34 structure parameters

层数	输出尺寸/(像素 × 像素)	网络结构
Conv1	112 × 112	7 × 7, 64
Conv2_x	56 × 56	$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix} \times 3$
Conv3_x	28 × 28	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix} \times 4$
Conv4_x	14 × 14	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix} \times 6$
Conv5_x	7 × 7	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix} \times 3$
输出层	1 × 1	平均池化层

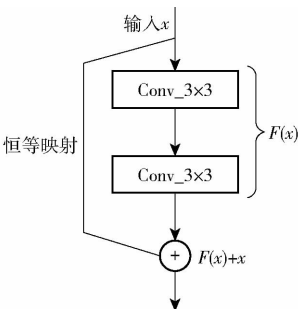


图 3 残差模块

Fig. 3 Residuals module

然而,本实验应用的鱼类不同饱腹程度数据集

具有鱼体目标较小,鱼群分布相对集中的特点,面对该类型的数据集,ResNet-34 存在感兴趣区域分散的问题,且由于该模型网络层数较多,参数量较大,会对后续在实际生产中的使用产生一定的阻碍。因此,本实验针对 ResNet-34 对鱼群不同饱腹程度分类过程中存在的不足,以 ResNet-34 作为网络基本骨架,对其进行一系列的改进。

2.2 模型改进

2.2.1 坐标注意力机制

注意力机制是人工神经网络中一种模仿人的认知注意力的技术。这种机制可以自动地根据特征矩阵中不同部分的重要性分配不同的权重,从而使得网络更加关注于重要的局部特征。目前注意力机制主要包括 Squeeze-and-excitation (SE)^[25] 注意力机制和 CBAM^[26] 注意力机制,但是这两种注意力机制仍然存在一些不足。SE 机制只注重于内部的通道信息而忽视位置信息;而位置信息是视觉任务获取物体结构的关键。CBAM 虽然尝试引入位置信息,但是其使用的全局池化的方式只能获取局部的位置,而无法获取长范围依赖的信息。由于无法获得准确而有效的位置信息,会影响权重的分配,从而影响到模型最终的准确率。

为了弥补以上缺陷,本文所使用的坐标注意力机制^[27] (Coordinate attention, CA) 将位置信息嵌入到通道注意力中,使得网络能够在较大的区域中进行注意力操作,同时也不会增大计算开销。为了解决二维全局池化所引起的位置信息损失问题,坐标注意力机制分别沿着垂直和水平方向利用两个一维全局池化操作将输入特征矩阵 ($C \times H \times W$) 池化为两个不同尺寸的特征矩阵 ($C \times H \times 1, C \times 1 \times W$), 公式为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h,i) \tag{1}$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j,w) \tag{2}$$

式中 W ——矩阵宽度 H ——矩阵高度
 $x_c(h,i)$ ——在第 c 个通道上,矩阵中位于第 h 行、第 i 列的元素
 $x_c(j,w)$ ——在第 c 个通道上,矩阵中位于第 j 行、第 w 列的元素

然后将垂直方向进行池化操作得到的矩阵 ($C \times 1 \times W$) 进行维度移位变成 ($C \times W \times 1$),然后将维度移位得到的矩阵与在水平方向进行池化操作得到的矩阵在垂直方向上进行拼接,矩阵结构变为 $C \times (H + W) \times 1$,使用 1×1 大小的卷积核对其进行卷积操作,进行标准化处理 (Batch normalization)^[28]

后将结果输入到 ReLU 激活函数^[29],公式为

$$f = \delta(F_1([z^h, z^w])) \tag{3}$$

式中 z^h ——在水平方向上进行一维池化后得到的矩阵
 z^w ——在垂直方向上进行一维池化后得到的矩阵

F_1 ——拼接函数 δ ——ReLU 激活函数

将得到的矩阵在垂直方向上进行分离,得到 2 个矩阵 ($C \times W \times 1, C \times H \times 1$),将矩阵 $C \times W \times 1$ 进行维度移位操作得到矩阵 $C \times 1 \times W$,对两个矩阵 $C \times H \times 1, C \times 1 \times W$ 分别进行卷积后进行升维操作得到矩阵 $C \times H \times W, C \times H \times W$,最后将这两个矩阵与恒等映射来的原矩阵进行乘法操作,最终得到结果矩阵,即

$$y_c(i,j) = x_c(i,j)g_c^h(i)g_c^w(j) \tag{4}$$

式中 $x_c(i,j)$ ——原矩阵第 c 个通道上,矩阵中位于第 i 行、第 j 列的元素
 $g_c^h(i)$ ——水平方向上进行一维池化后得到的新矩阵, $g_c^h(i)$ 为该新矩阵第 c 个通道的第 i 个元素
 $g_c^w(j)$ ——垂直方向上进行一维池化后得到的新矩阵, $g_c^w(j)$ 为该新矩阵第 c 个通道的第 j 个元素
 $y_c(i,j)$ ——结果矩阵第 c 个通道上,矩阵中位于第 i 行、第 j 列的元素

通过上述操作,得到两个独立的方向感知特征映射,并将其编码为两种注意力映射,以获取输入特征映射在不同空间方向上的长距离依赖关系,从而将位置信息保存到生成的注意力图中,最后通过乘法操作将两个注意力图应用于特征图当中,保证在较大的范围内完成对特征的有效提取。

如图 4a 所示,本实验将坐标注意力模块添加进 ResNet 的残差模块中 (3 ~ 18 层的 BasicBlock),总计 16 层。此外,CA 的结果如图 4b 所示。

2.2.2 深度可分离卷积

本文所使用的骨干网络 ResNet34 通过不断堆叠网络层数,并且通过使用 shortconnection 结构避免了梯度爆炸或消失等问题,真正做到了网络越深,提取特征能力越强,最终的准确率越高。但是随着网络层数的增加也带来了模型参数量过大的问题,给后续模型在移动端的部署造成阻碍。因此,需要在尽量不影响模型准确率的前提下大大减少模型的参数量,从而为后续模型在移动端部署提供可能。本文提出使用深度可分离卷积 (Depthwise separable convolutions) 的方式来替换 ResNet 中 Basicblock 的第 2 个传统卷积,如图 5 所示。

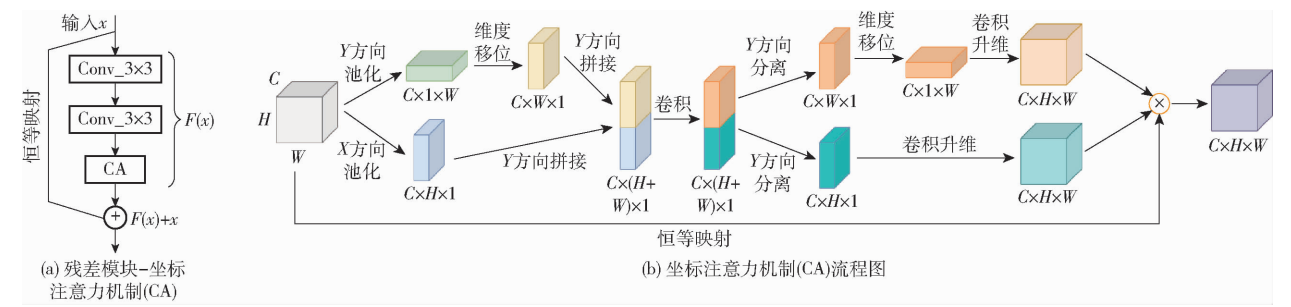


图 4 残差模块-坐标注意力和坐标注意力机制流程图
Fig. 4 Residual module - CA and coordinate attention process

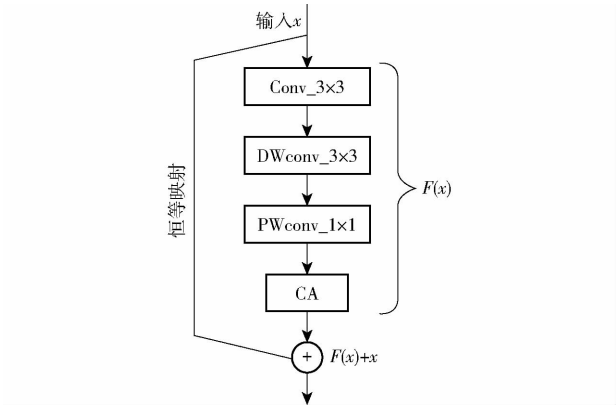


图 5 深度可分离卷积-坐标注意力机制
Fig. 5 Depthwise separable convolutions - CA

从图 5 中可以看出深度可分离卷积由深度卷积 (Depthwise convolution) 和逐点卷积 (Pointwise convolution) 两部分组成,如图 6 所示。从图 6a 中可以看出,深度卷积所使用的卷积核个数与输入特征矩阵通道数量一致,并且卷积核的通道数为 1,此操作可以减少大量参数,但是深度卷积只是对输入层的每个通道独立进行卷积运算,割裂了输入特征图不同通道在相同空间位置上的联系,因此后续还需要通过逐点卷积来将新生成的输出特征矩阵作为输入层,最终实现特征图在深度方向的加权操作。其中从图 6b 中可以看出逐点卷积与传统卷积类似,不同之处在于逐点卷积所使用的卷积核大小均为 1×1 。

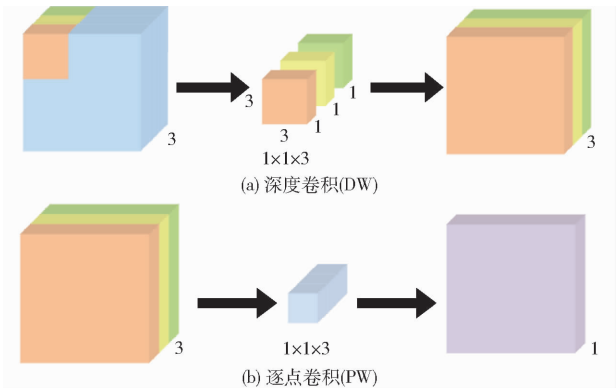


图 6 深度卷积和逐点卷积

Fig. 6 Depthwise convolution and pointwise convolution

传统卷积与深度可分离卷积参数量之比为 $1/N$,其中 N 为卷积核数量。从化简结果可以看出,本文所使用的深度可分离卷积相较于传统卷积可以减少模型所使用的参数。通过深度卷积与逐点卷积的组合运算,做到了在牺牲少量准确率的基础上,极大地减少了模型的参数量。

3 实验与结果分析

3.1 实验环境

实验均在 Windows 10 64 位系统环境下运行,采用能够支持 GPU 加速和动态神经网络的 Pytorch1.10 版本的深度学习开源框架,配合 CUDA 11.0 进行训练。计算机搭载的处理器为 IntelCorei5 - 7300HQCPU,2.50 GHz 4 核,内存为 8 GB,显卡为 GTX1050Ti。

3.2 参数设置

实验过程中将数据集按 7:3 划分为训练集和测试集,分别用于模型的训练与测试。所有模型的全连接层修改为 5,代表 5 种不同的鱼类饱腹程度。鱼类饱腹程度图像尺寸均设置为 224 像素 $\times$ 224 像素,批量大小 (Batch size) 则设置为 4,模型迭代次数共计 30 周期。学习率设为 0.0001,模型采用 Adam 优化器,其参数设置为默认。

3.3 消融实验结果分析

为了证明本文实验改进的有效性,在 ResNet - 34 相同的位置分别添加了 SE 注意力模块、CBAM 注意力模块,以及本文最终所使用的 CA 注意力模块,并应用到测试集上,最终得到 4 种模型的识别准确率分别为 90.0%、91.9%、91.7%、93.8%,结果表明,相比其他 3 种类型,添加了 CA 注意力模块的 ResNet - 34 能够更精确地实现对鱼类不同饱腹程度的分类。

为了降低模型的参数量,实验将加入 CA 模块的 ResNet - 34 模型中的第 2 个普通卷积替换成了深度可分离卷积与点卷积的组合,实验结果表明,本模型的最终准确率达到 93.4%,相较于未替换的模

型准确率只降低 0.4 个百分点,而参数量相较于之前减少 46.7%。

图 7 为原模型与改进模型的准确率变化曲线。3 种模型的参数量与准确率如表 3 所示。

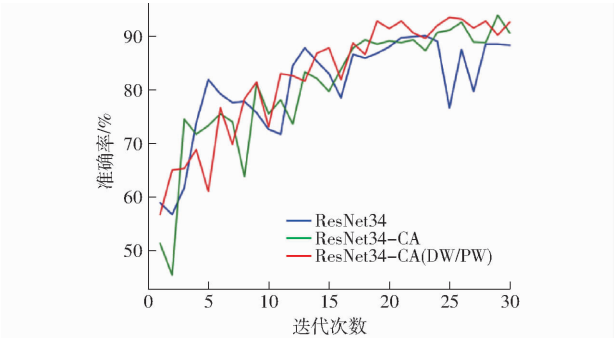


图 7 原模型与改进模型的准确率变化曲线
Fig. 7 Accuracy change curves of original model and improved model

表 3 基于 ResNet34 的消融实验结果
Tab. 3 Results of ResNet34-based ablation tests

模型	参数量	准确率/%
ResNet34	2.18×10^7	90.0
ResNet34-CA	2.19×10^7	93.8
ResNet34-CA(DW/PW)	1.19×10^7	93.4

从图 7 可以看出,改进后的模型可以达到最高的准确率,在训练到第 20 个周期时模型逐渐开始收敛,在使用坐标注意力机制的情况下,即使是将传统卷积替换为深度可分离卷积,准确率仅下降 0.4 个百分点,且从图中可以看出,模型也能做到稳定收敛。然而 ResNet34 模型在训练后期会出现较大的波动,且预测准确率只有 90%。综上,通过以上改进使模型的预测准确率以及参数内存方面取得了良好的平衡,为后续模型在移动端的部署提供了可能。

3.4 模型比较

为了验证改进模型与其他传统 CNN 模型的优势,将改进后的模型与 AlexNet^[30]、VGG16^[31]、MobileNet-v2^[32]、GoogLeNet^[33] 进行比较,为了保证实验的公平性,均采用相同的网络参数进行训练。

训练后结果如表 4 所示,从表中可以看出,使用了深度可分离卷积的 ResNet34-CA 模型在测试集

表 4 各网络模型分类对比
Tab. 4 Comparison of classification of each network model

模型	准确率	精确率	召回率	F1 值
ResNet34-CA(DW/PW)	93.4	93.8	93.2	93.4
MobileNet-v2	88.4	88.7	87.3	87.3
VGG16	89.6	90.2	89.4	89.4
GoogLeNet	89.2	88.9	87.7	87.6
AlexNet	85.8	85.2	84.0	84.0

中的各项数据均达到了最优水平。相比之下, MobileNet-v2 和 GoogLeNet 效果表现较差, AlexNet 模型表现最差。VGG16 模型的准确率达到 89.6%, 与实验所使用的原模型 ResNet34 相比只低 0.4 个百分点,但是 VGG16 的参数量 (4.25×10^7) 是 ResNet34 的两倍左右。

通过与其他 CNN 模型进行比较,本文提出的深度可分离卷积 ResNet34-CA 模型在鱼群不同饱腹程度数据集中取得了最优成绩。通过在原模型中添加坐标注意力机制,增加了模型对全局重要信息的关注度,从而提高了算法性能。并且从图 8 中可以看出,改进之后的模型相较于其他模型更加稳定,其他模型均存在一定程度的波动。

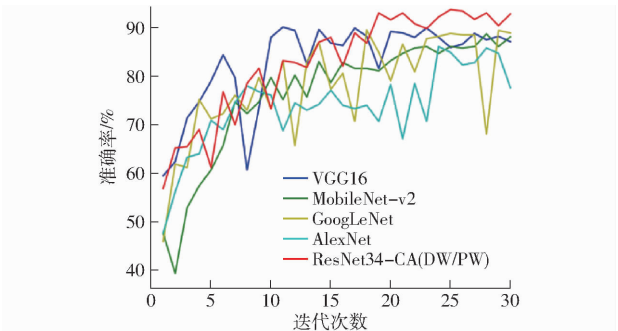


图 8 5 种 CNN 架构的分类准确率对比
Fig. 8 Comparison of classification accuracy of 5 CNN architectures

3.5 Grad-CAM 特征可视化结果

为了更加清晰展示本文模型的效果,根据 Grad-CAM^[34] 方法,绘制了使用 3 种不同注意力机制的网络最后一层可视化热力图和 ResNet34 原模型最后一层网络热力图,如图 9 所示。原始图像为从测试集中随机选择得到。从图中可以看出,加入了 3 种不同注意力机制后模型所关注的区域相较于原模型更加集中于鱼群聚集的区域,但是使用 SE 注意力机制与 CBAM 注意力机制的模型提取效果相较于加入 CA 注意力机制的模型相对较差,对于鱼类摄食特征关注范围较小,且无法完全注意到鱼群聚集区域。使用 CA 注意力机制的模型所关注的范围更大,并且可以精确定位到鱼体聚集的区域。这也是

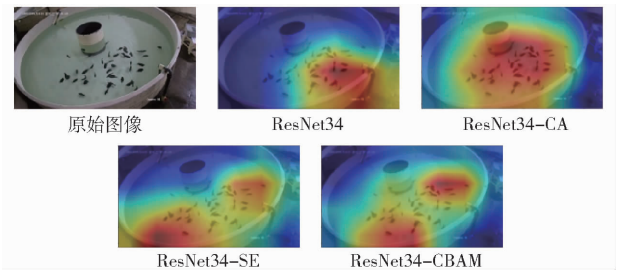


图 9 不同注意力机制可视化热力图
Fig. 9 Visual heat map of different attention mechanisms

在对 2.2.1 节所提到 CA 注意力机制相较于其他注意力机制所具有优势的合理解释,体现出对原模型改动的合理性。

4 结论

(1)本文模型(ResNet34-CA(DW/PW))对于鱼群不同饱腹程度识别具有更强的分类性能,使用了坐标注意力机制使得模型可以关注到更大范围内的有效信息,最终测试集准确率达到 93.4%,相比于 AlexNet、GoogLeNet、VGG16、MobileNet-v2、ResNet34,本文模型在测试集上分别提高 7.6、4.2、3.8、5.0、3.4 个百分点,并且在精确率、召回率和 F1 值 3 项指标中,本文模型也优于其他

模型。

(2)通过使用深度可分离卷积来替换传统卷积,ResNet34-CA(DW/PW)相较于 ResNet34-CA 模型参数量减少 46.7%,而准确率仅降低 0.4 个百分点,综合模型性能与模型内存占用量,ResNet34-CA(DW/PW)更优。

(3)所提出的方法避免了声学传感器价格昂贵且易损坏的问题,通过视频传感器获取数据也不会对水环境造成入侵。并且本文使用的深度学习方法,克服了需要对鱼类行为进行量化等问题,最终通过对 ResNet34 模型进行的一系列改进,使得模型在准确率和参数量之间达到了平衡,为后续在实际养殖环境中应用提供了可能性。

参 考 文 献

[1] 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.

[2] ATOUM Y, SRIVASTAVA S, LIU X. Automatic feeding control for dense aquaculture fish tanks[J]. IEEE Signal Processing Letters, 2015, 22(8): 1089-1093.

[3] 马长震,谌志新,徐志强,等. 基于超声探测技术的深水网箱剩余饵料监测系统[J]. 微计算机信息, 2012, 28(4): 39-40.

[4] 石韵. 饲料系数及价格对水产养殖经济的影响[J]. 中国饲料, 2020(8): 84-87.

[5] 袁超,朱瑞金. 基于 KPCA 的多特征融合的支持向量机鱼群摄食行为检测研究[J]. 水产养殖, 2020, 41(12): 17-21.

[6] PAPADAKIS V, PAPADAKIS I, LAMPRIANIDOU F, et al. A computer-vision system and methodology for the analysis of fish behavior[J]. Aquacultural Engineering, 2012, 46: 53-59.

[7] 胡金有,王靖杰,张小栓,等. 水产养殖信息化关键技术研究现状与趋势[J]. 农业机械学报, 2015, 46(7): 251-263.

[8] HU Jinyou, WANG Jingjie, ZHANG Xiaoshuan, et al. Status and trends of research on key technologies for aquaculture informatization[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(7): 251-263. (in Chinese)

[9] LI D, WANG Z, WU S, et al. Automatic recognition methods of fish feeding behavior in aquaculture: a review[J]. Aquaculture, 2020, 528: 735508.

[10] ZHOU C, XU D, LIN K, et al. Intelligent feeding control methods in aquaculture with an emphasis on fish: a review[J]. Reviews in Aquaculture, 2018, 10(4): 975-993.

[11] ZHANG H, WEI Q, KANG M. Measurement of swimming pattern and body length of cultured Chinese sturgeon by use of imaging sonar[J]. Aquaculture, 2014, 434: 184-187.

[12] 吴伟,范立民. 水产养殖环境的污染及其控制对策[J]. 中国农业科技导报, 2014(2): 9.

[13] 杜海明. 投喂策略对鳃鲶鱼摄食、生长及体成分的影响[D]. 武汉:华中农业大学, 2007.

[14] 贾成功,张学良,陈俊华,等. 基于鱼群摄食规律的投饵系统研究[J]. 机械工程师, 2017(8): 22-25, 28.

[15] YE Z, ZHAO J, HAN Z, et al. Behavioral characteristics and statistics-based imaging techniques in the assessment and optimization of tilapia feeding in a recirculating aquaculture system[J]. Transactions of the ASABE, 2016, 59(1): 345-355.

[16] ZHAO J, BAO W, ZHANG F, et al. Assessing appetite of the swimming fish based on spontaneous collective behaviors in a recirculating aquaculture system[J]. Aquacultural Engineering, 2017, 78: 196-204.

[17] 黄志涛,何佳,宋协法. 基于鱼体运动特征和图像纹理特征鱼类摄食行为识别与量化[J]. 中国海洋大学学报(自然科学版), 2022, 52(1): 32-41.

[18] 荆丹翔,韩军,王英杰,等. 基于成像声呐的鱼类三维空间分布[J]. 水产学报, 2018, 42(6): 996-1005.

[19] MALLEKH R, LAGARDERE P, ENEAU P, et al. An acoustic detector of turbot feeding activity[J]. Aquaculture, 2003, 221(1-4): 481-489.

[20] KOLAREVIC J, HANSEN A, ESPMARK A, et al. The use of acoustic acceleration transmitter tags for monitoring of Atlantic salmon swimming activity in recirculating aquaculture systems (RAS)[J]. Aquacultural Engineering, 2016, 72-73: 30-39.

[21] 陈刚,陈卫忠. 渔业资源评估中声学方法的应用[J]. 上海水产大学学报, 2003(1): 40-44.

[22] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature (London), 2015, 521(7553): 436-444.

[23] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[24] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.