

# 基于 MobileViT - CBAM - BiLSTM 的开放式养殖环境 鱼群摄食强度分类模型

徐立鸿<sup>1,2</sup> 黄志尊<sup>1</sup> 龙伟<sup>1</sup> 蒋林华<sup>1</sup> 童欣<sup>1</sup>

(1. 湖州师范学院信息工程学院, 湖州 313000; 2. 同济大学电子与信息工程学院, 上海 201804)

**摘要:** 鱼群摄食的精准投喂技术是实现智慧化水产养殖的一项关键技术。大多数精准投喂模型都是基于水质较清晰的室内养殖池,不太适用于开放式养殖环境。本研究通过水上视角采集构建了一套开放式池塘数据集,并对数据集进行数据增强增加其多样性,然后在轻量化神经网络 MobileViT 基础上,将 CBAM 注意力模块与 MV2 模块结合设计了 CBAM - MV2 模块,并嵌入 BiLSTM 循环神经网络用于识别分类,提出改进的 MobileViT - CBAM - BiLSTM 模型,提高了模型预测能力、鲁棒性和泛化性能,实现了鱼群摄食行为的三分类。实验结果显示,改进后 MobileViT 在采集的视频帧数据集上明显优于改进前的 MobileViT,准确率 98.61%,宏 F1 值达 98.79%,相对于原始 MobileViT 准确率提高 6.33 个百分点,宏 F1 值提高 6.75 个百分点。

**关键词:** 鱼群摄食强度分类模型; 精准投喂; MobileViT; BiLSTM; CBAM

中图分类号: S951.2 文献标识码: A 文章编号: 1000-1298(2024)11-0147-07

OSID:



## Classification Model of Fish Feeding Intensity Based on MobileViT - CBAM - BiLSTM

XU Lihong<sup>1,2</sup> HUANG Zhizun<sup>1</sup> LONG Wei<sup>1</sup> JIANG Linhua<sup>1</sup> TONG Xin<sup>1</sup>

(1. School of Information Engineering, Huzhou University, Huzhou 313000, China

2. College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** Precise feeding technology for fish ingestion is a key technology to achieve intelligent aquaculture. However, most of the precise feeding model is based on indoor aquaculture ponds with clear water quality, which are not suitable for outdoor open farming environments. In view of the actual situation, a set of detailed open pond dataset through water perspective acquisition was constructed, and the dataset was augmented to increase its diversity, and then the BiLSTM bidirectional recurrent neural network was embedded on the basis of the lightweight neural network MobileViT, so as to improve the memory ability of the model for video sequence data in a long period of time, and the CBAM attention module was combined with the MV2 module to design the CBAM - MV2 module, and then the CBAM - MV2 module was added to different layers of the model for experiments to obtain the most reasonable improvement scheme. Finally, an improved MobileViT - CBAM - BiLSTM fish feeding behavior classification model was proposed, which improved the prediction ability, robustness and generalization performance of the model, and realized the three classification of fish feeding behavior. The experimental results showed that the improved MobileViT was significantly better than previous in the collected video frame dataset, with an accuracy of 98.61%, 98.79% for Macro-F1, which was 6.33 percentage points for accuracy, 6.75 percentage points for Macro-F1 compared with the original MobileViT.

**Key words:** classification model of fish feeding intensity; precision feeding; MobileViT; BiLSTM; CBAM

收稿日期: 2024-01-08 修回日期: 2024-04-22

基金项目: 国家自然科学基金项目(62373286,62175037)和湖州市重点研发计划农业“双强”专项(2022ZD2060)

作者简介: 徐立鸿(1960—),男,教授,博士生导师,主要从事水产养殖机器视觉技术等研究,E-mail: xulihong@tongji.edu.cn

通信作者: 龙伟(1978—),男,讲师,博士,主要从事图像处理、人工智能和水产养殖中的视觉技术等研究,E-mail: lw@zhu.edu.cn

## 0 引言

近年来海洋捕捞比例持续增加,为缓解捕捞压力,水产养殖产量在持续增加<sup>[1]</sup>。大多数养殖户也开始使用自动饲喂机进行投喂<sup>[2]</sup>。但使用自动喂料机可能造成喂料过量或喂料不足的情况<sup>[3]</sup>。摄食不足易造成鱼体营养不良,且降低养殖经济收益。相反,过度饲养会增加饲养成本,并影响鱼类健康生长,污染养殖环境<sup>[4-6]</sup>。而高精度识别鱼类摄食行为可以有效地指导摄食过程并根据饥饿程度提供最佳投喂量<sup>[7-8]</sup>。因此,科学准确地识别鱼群食欲程度是实现水产养殖现代化、工业化的关键一步。

文献[9-11]表明,计算机视觉可以用于估计鱼的食欲以实现精确喂养。但相关研究<sup>[12-15]</sup>大多都基于室内循环水养殖池,不太适用于开放式户外养殖池。

UBINA等<sup>[16]</sup>提出一种两阶段的方法来改进鱼的摄食强度评估:首先应用光流神经网络(Optical flow, OF)生成光流帧,然后将光流帧输入到3D卷积神经网络(3D convolutional neural networks, 3DCNN)中进行鱼的摄食强度评估。唐宸等<sup>[17]</sup>利用光流法进行特征提取,实现水下鱼群摄食状态三分类及水上鱼群摄食状态二分类。徐立鸿等<sup>[18]</sup>基于深度学习模型LRCN(Long-term recurrent convolutional networks)进行改进,将SE(Squeeze-and-excitation networks)注意力模块融入LRCN,提高模型特征拾取能力用于鱼群摄食行为分类,并实现水上摄食状态的三分类。

声波遥测技术在传统观测技术之上进行了补充,提供了监测鱼类的机会,而不必直接与它们互动或将它们与其他个体隔离。ZENG等<sup>[19]</sup>提出一种基于声信号和注意力机制的音频频谱Swin Transformer模型,该模型可以将鱼类的摄食强度分为4个等级:强、中、弱和无。但用于采集声信号的声呐设备价格昂贵,不适用于大部分养殖环境。

MobileViT<sup>[20]</sup>是一款轻量化的可用于移动设备的通用视觉识别模型,其结合CNN和ViT的优势,有着低延迟、高精度等特点。BiLSTM<sup>[21]</sup>(Bidirectional long short-term memory)是一种双向循环神经网络,可以有效地使用过去和未来的输入特征,使网络更能捕捉到准确的信息,应对更加复杂的情况,提高鲁棒性,在各种数据及任务中都表现出其优异的成绩。CBAM<sup>[22]</sup>(Convolutional block attention module)是一种简单而有效的前馈卷积神经网络注意力模块。结合卷积和注意力机制,可以从空间和通

道两方面上对图像进行关注,提高模型泛化能力和性能。

针对现实生产环境情况,本文提出一种基于改进MobileViT的鱼群摄食行为分类模型,与CBAM注意力模块和BiLSTM双向循环神经网络相结合,搭建MobileViT-CBAM-BiLSTM模型,通过采集水上视频帧数据集,并经过数据增强扩充数据集,增加样本多样性,实现鱼群摄食行为三分类,为开放式养殖池精准投喂提供技术支持。

## 1 材料与方法

### 1.1 试验材料与环境

本文测试对象为草鱼、鲫鱼混合鱼群,这些鱼类性活泼、游泳迅速、常成群觅食,具有较强的夺食性,可在养殖池塘水面上观察到较为明显的强度差异,有利于算法测试与评估。采集养殖基地为上海瑞钵水产有限公司,该基地位于上海市崇明区。数据采集池塘为3号池塘,长约50 m,深约2 m,宽约25 m,面积约为2 500 m<sup>2</sup>,并配备了一台广东省佛山市中渔科技有限公司生产的ZY150型投饵机。

池塘采取定时定量投饵方案,将09:00—16:00每隔1 h进行间歇式分组,组间采取少量多次投饵方案,每次投完饵料后间隔1 min,有利于对食饵情况进行判断。

### 1.2 视频采集方案

视频结合现场实际情况与研究实际情况,采用水上拍摄方法,自上而下地采集池塘中鱼群的摄食视频数据集,示意图如图1所示。采集该视频数据集的摄像机为海康威视DS-2CD3T47DWD-L型摄像机。主要拍摄中心区域为投饵机饵料抛撒区,选择不同时间段分别拍摄以取得样本的多样性。在光照条件中等或较弱时拍摄时长为15 min,在光照条件较强时拍摄时长延长一倍以对抗过曝等不可抗条件。样本拍摄初始分辨率为1 920像素×1 440像素,帧率为30 f/s。视频采集时长如表1所示。

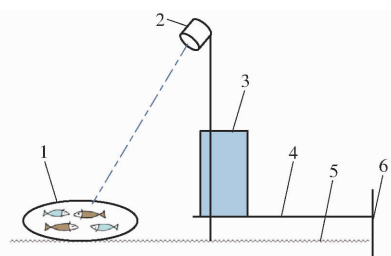


图1 视频采集平台结构示意图

Fig. 1 Schematic of video capture platform

1. 拍摄目标区域
2. 摄像机
3. 投饵机
4. 木板
5. 水面
6. 陆地

表 1 视频采集时长

Tab. 1 Video capture duration

天气	时间段	光照条件	采集时长/min
晴	09:00—10:00	中等	15
	12:00—13:00	较强	30
	15:00—16:00	中等	15
阴	09:00—10:00	较弱	15
	12:00—13:00	中等	15
	15:00—16:00	较弱	15

### 1.3 数据集制作

根据文献及养殖专家的经验将视频数据集分成无摄食、弱摄食与强摄食 3 类。无摄食时鱼群对投喂的饵料无明显反应,水面平静且几乎无上浮鱼群;弱摄食时水面有少量清晰可见的鱼群且伴随鱼群主动进食动作,水面有较少水花;强摄食时水面有大量清晰可见的鱼群且伴随鱼群争抢进食动作,水面水花明显。视频帧图像如图 2 所示。

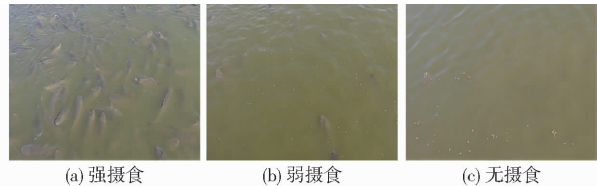


图 2 视频帧数据集示例图像

Fig. 2 Video frame data set

将分好的长视频数据集裁剪成利于模型训练的多个 2 s (60 帧) 短视频数据集,并将尺寸裁剪为 416 像素 × 416 像素。根据现场情况,可以适当调整视频长度、尺寸(像素)以更好匹配实际情况。

数据集中强摄食、弱摄食、无摄食的比例约为 1:1:1,将数据集按比例 9:1 划分为训练集和测试集,再将训练集按照比例 9:2 划分为训练集和验证集。为增加模型鲁棒性和泛化性,选择不同天气下的训练集进行数据增强,主要包括翻转、归一化、添加高斯噪声等常见方式。设置高斯噪声均值为 0,标准差为 0.001。除此之外,为减少光照等自然条件对模型分类性能的影响,以阴天和晴天为例,对训练集数据添加图像色彩抖动,色彩抖动方法主要包括随机调整图像的对比度、锐度、亮度和饱和度等。色彩抖动效果如图 3 所示。通过对数据集增加色彩抖动效果,提高模型泛化能力。增强后总训练集数量为 889,验证集数量为 225,测试集数量为 110,数据集划分如表 2 所示。

## 2 鱼群摄食行为分类改进 MobileViT 模型

### 2.1 MobileViT

MobileViT 是一款用于移动设备的轻型通用视

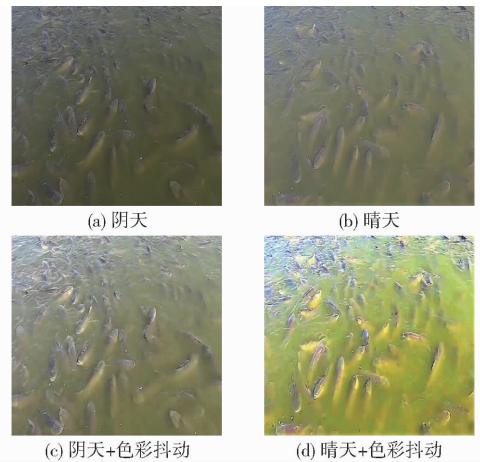


图 3 不同天气图像添加色彩抖动效果

Fig. 3 Adding color shake to different weather images

表 2 数据集视频数量

Tab. 2 Data set number

类型	训练集	验证集	测试集
强摄食	308	78	37
弱摄食	286	72	36
无摄食	295	75	37
总计	889	225	110

觉识别模型,常用于分类任务中,其网络结构如图 4(图中 MV2 表示 MobileNetV2 块<sup>[23]</sup>,其中执行下采样的块被标记为 ↓2,下同)所示。其引入 MobileViT block,将卷积中的局部处理替换为使用 transformer 的全局处理,使得 MobileViT 既有 CNN 的参数少、速度快的特点,又有 ViT 关注全局的优势,为有轻量、低延迟、高精度需求的任务提供了可能性。

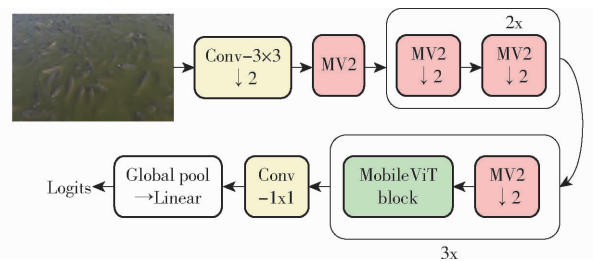


图 4 MobileViT 结构图

Fig. 4 Structure of MobileViT

在 MobileViT 结构中,输入图像先经过一个标准的  $3 \times 3$  卷积层进行卷积,然后经过多个 MV2 和 MobileViT block 的持续下采样和池化提取特征,再连接一个  $1 \times 1$  的卷积层,最终通过全连接输出分类结果。

### 2.2 CBAM

CBAM 是一种简单而有效的前馈卷积神经网络注意模块。给定一个中间特征映射,CBAM 沿着通道和空间两个独立的维度依次推断注意力映射,然后将注意力映射乘以输入特征映射以进行

自适应特征细化。由于结合了通道和空间注意力, CBAM 能降低对噪声及不相关杂波的关注度。它可以无缝地集成到任何 CNN 架构中, 与基础 CNN 一起进行端到端训练。CBAM 示意图如图 5 所示。

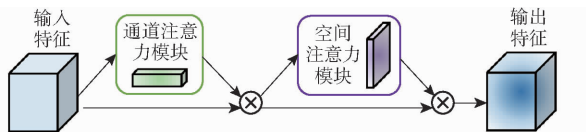


图 5 CBAM 模块示意图

Fig. 5 Structure of CBAM block

以中间特征图  $F \in \mathbf{R}^{C \times H \times W}$  为输入, CBAM 依次推导出二维通道注意力图  $M_c \in \mathbf{R}^{C \times 1 \times 1}$  和二维空间注意力图  $M_s \in \mathbf{R}^{1 \times H \times W}$ 。整个注意力过程为

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

其中,  $\otimes$  表示乘法。在乘法过程中, 特征值相应地被通道注意力和空间注意力计算。

### 2.3 BiLSTM

在时间序列任务中, 访问给定时间内过去和未来的输入特征会让任务变得更有效, 因此可以使用双向 LSTM 网络来实现这一目的<sup>[21]</sup>。如图 6 所示, 与 LSTM 相比, BiLSTM 可以有效地利用特定时间框架内的过去特征和未来特征, 其具有更高的鲁棒性, 并且受工程特征去除的影响较小<sup>[24]</sup>。

### 2.4 改进的 MobileViT 模型

以 MobileViT 为基础, 提出一种基于视频数据注意力融合的 MobileViT - CBAM - BiLSTM 网络。鱼群摄食行为分类是根据一段时间内鱼群摄食特征

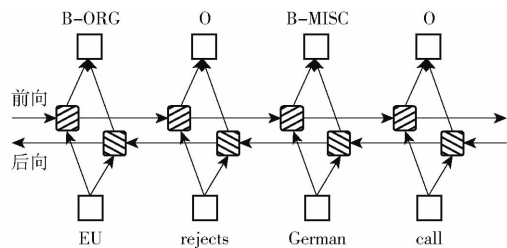


图 6 BiLSTM 模块示意图

Fig. 6 Structure of BiLSTM

来区分鱼群摄食强度的任务, 其视频数据中包含很多视频帧, 处理这些视频帧所付出的计算代价很大, 也会降低识别与类别相关的视频帧的性能。因此, 本文在 MobileViT 的残差模块中引入 CBAM 注意力模块增强通道和空间的表征, 来识别与目标语义直接相关的最有判别力的时空特征。其次将双向时序循环网络 BiLSTM 与 MobileViT 进行结合, 利用循环网络对序列数据的优势强化模型对视频数据的处理能力, 这于开放式池塘环境中识别鱼群摄食行为有较大的增益。最后采用不同的训练策略进行训练比较, 优化模型性能。

MobileViT - CBAM - BiLSTM 具有轻量化、高精度、高鲁棒性等特点。模型结构如图 7 所示。其中 MobileViT\_CBAM 负责特征提取, BiLSTM 进行类别识别。输入的视频帧数据集首先经过一个标准的  $3 \times 3$  卷积层进行卷积, 然后经过多个 MV2 模块与 CBAM - MV2 模块和 MobileViT block 模块的下采样池化, 将池化后的结果输入 BiLSTM 双层循环网络中进行预测分类, 最终预测出鱼群摄食类别。

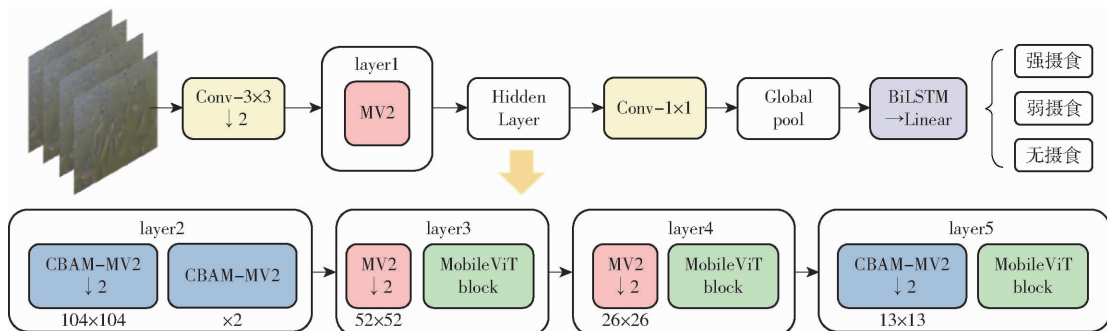


图 7 MobileViT - CBAM - BiLSTM 结构图

Fig. 7 Structure of MobileViT - CBAM - BiLSTM

其中 CBAM - MV2 模块基于 MV2 模块改进而来, 添加 CBAM 模块提高其提取特征的能力。其由一个顺序的  $1 \times 1$  卷积、 $3 \times 3$  卷积、CBAM 注意力模块和  $1 \times 1$  卷积组成, 结构如图 8 所示。

### 2.5 评价指标

为综合反映模型整体性能, 选择准确率 (Accuracy)、微 F1 值、宏 F1 值和加权 F1 值<sup>[25]</sup> 4 个

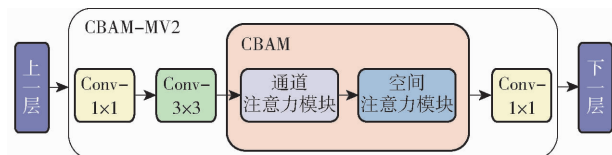


图 8 CBAM - MV2 结构图

Fig. 8 Structure of CBAM - MV2

指标作为评价指标。



### 3 结果与分析

#### 3.1 实验配置

实验硬件环境为 Intel(R) Xeon(R) CPU E5 - 2698 v4 @ 2.20 GHz、80 GB 内存、NVIDIA Corporation GA100 [A100 PCIe 40 GB] 显卡。软件环境基于 Linux 操作系统, 编程语言为 Python 3.8, CUDA 版本为 12.2, 深度学习框架使用 Pytorch 2.1.0。

#### 3.2 循环网络

为选择更加契合实验条件的循环网络, 以 MobileViT 作为基准模型, 选择添加 LSTM<sup>[26]</sup>、传统 RNN<sup>[27]</sup>、GRU<sup>[28]</sup> 以及 BiLSTM 作为本实验的循环网络结构进行实验对比, 结果如表 3 所示。从表 3 可以看到, BiLSTM 相较于 LSTM、传统 RNN、GRU 更契合本实验, 准确率达 97.92%, 宏 F1 值达 98.03%, 加权 F1 值达 97.90%。

表 3 不同循环网络模型性能

Tab. 3 Performance of models under different cyclic network structures %

网络模型	准确率	宏 F1 值	微 F1 值	加权 F1 值
RNN	93.06	92.83	93.06	93.32
GRU	96.53	96.38	96.53	96.55
LSTM	93.75	93.27	93.75	93.82
BiLSTM	97.92	98.03	97.92	97.90

#### 3.3 注意力模块

为验证在不同残差模块中添加 CBAM 注意力模块对网络的影响, 基于 MobileViT - BiLSTM 网络在第 2 层、第 3 层、第 4 层、第 5 层中添加 CBAM 模块, 其中“√”表示添加 CBAM 模块, “-”表示不添加, 使用原始模块, 结果如表 4 所示。

表 4 CBAM 在模型不同层中效果

Tab. 4 Effect of CBAM on model at different layers %

第 2 层	第 3 层	第 4 层	第 5 层	准确率	加权 F1 值
√	-	-	-	97.50	97.49
-	√	-	-	96.25	96.26
-	-	√	-	95.63	95.68
-	-	-	√	94.37	94.36
√	√	-	-	96.62	96.65
√	-	√	-	97.31	97.21
√	-	-	√	98.61	98.61
√	√	√	√	96.88	96.88

由表 4 可知, 单层 CBAM 在第 2 层、第 3 层、第 4 层、第 5 层中的准确率分别为 97.50%、96.25%、95.63%、94.37%; 在双层 CBAM 中同时第 2 层和

第 5 层添加的准确率达到最高, 为 98.61%; 在 4 层都添加的情况下模型因变得复杂后受到过拟合的影响准确率反而有所下降。因此, 考虑到准确性和网络参数的数量, 本研究选择在第 2 层和第 5 层添加注意力模块进行下一步改进。

SE 模块和 CBAM 模块都是深度学习中常用的注意力模块, SE 是具有通道维度注意力, CBAM 是具有通道和空间双通道维度注意力, 为验证这两个模块对本实验的影响程度, 第 2 层和第 5 层同时分别添加 SE 模块和 CBAM 模块进行实验, 结果如表 5 所示。从表 5 看出, CBAM 模块比 SE 模块更契合本实验, 各方面综合表现也更好。相较于 SE 模块, 添加 CBAM 模块的网络准确率提高 1.73 个百分点, 宏 F1 值提高 1.74 个百分点, 加权 F1 值提高 1.76 个百分点。

表 5 不同注意力模块模型性能

Tab. 5 Performance of models with different attention module structures %

注意力模块	准确率	宏 F1 值	微 F1 值	加权 F1 值
SE	96.88	97.05	96.88	96.85
CBAM	98.61	98.79	98.61	98.61

#### 3.4 MobileViT - CBAM - BiLSTM 模型结果分析

混淆矩阵是分类任务中常用的可视化工具, 能较为清晰地看到各个类别在模型中的分类情况。因此本研究在测试集上制作了三类别的混淆矩阵, 结果如图 9 所示。

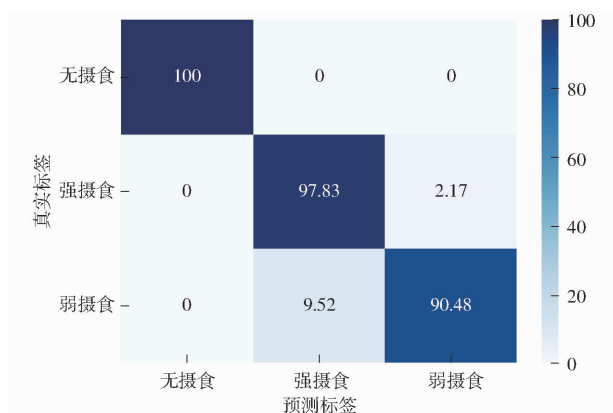


图 9 测试集混淆矩阵

Fig. 9 Test set confusion matrix

在混淆矩阵中, 颜色的深浅直观地反映了准确率大小。观察图 9 可以清晰地看到模型在测试集上对于无摄食和强摄食的识别表现很好, 其准确率分别高达 100% 和 97.83%。而弱摄食准确率仅为 90.48%, 相对较低。深入分析这一现象, 发现无摄食和强摄食行为特征在视频中通常表现得较为鲜明, 使得模型能够较为准确地捕捉并识别。相比之

下,弱摄食特征则相对模糊,可能介于无摄食和强摄食之间,这导致模型在识别时存在一定难度,有时会将弱摄食误判为强摄食。并且一段摄食视频中可能夹杂着多种摄食强度,在一定程度上会降低模型准确率。

将 MobileViT - CBAM - BiLSTM 和改进前的 MobileViT 进行对比,结果如表 6 所示。同时选择徐立鸿等<sup>[18]</sup>同样应用于开放式养殖池鱼群摄食行为的算法模型 SE - LRCN 进行对比。由表 6 可知, MobileViT - CBAM - BiLSTM 相较于原始 MobileViT 性能有明显提升,准确率提高 6.33 个百分点,宏

F1 值提高 6.75 个百分点,加权 F1 值提高 6.31 个百分点。相较于 SE - LRCN,准确率提高 1.61 个百分点,加权 F1 值提升 3.39 个百分点。测试样本分类结果如图 10 所示。

表 6 不同模型实验结果

Tab. 6 Experimental results of each model				%
模型	准确率	宏 F1 值	微 F1 值	加权 F1 值
MobileViT	92.28	92.04	92.28	92.30
MobileViT - CBAM - BiLSTM	98.61	98.79	98.61	98.61
SE - LRCN <sup>[18]</sup>	97.00	94.80	97.00	95.22



图 10 测试样本分类结果示例

Fig. 10 Diagrams of classification results of test samples

实验结果表明:①得益于 BiLSTM 拥有前向及后向双向的巡视,相较于传统 RNN、GRU、LSTM 拥有更强的识别能力,在更加复杂的环境下拥有更高的准确度。②CBAM 是拥有空间和通道两个维度的注意力模块,其能提高模型对于重要特征的关注度,相比于单通道 SE 注意力模块拥有更好的目标特征注意能力,使模型表现出更好性能,并提高鲁棒性及泛化性能。③在强化网络提取特征能力时并非添加的模块数量越多效果越好,有时会因添加模块过多造成过拟合,从而导致精度下降。

## 4 结论

(1)提出一种 MobileViT - CBAM - BiLSTM 模型,通过结合视频空间信息和时序信息实现对鱼群的摄食行为分类。针对开放式养殖池的鱼群摄

食行为实现了强摄食、弱摄食和无摄食三分类,推动了鱼群精准投喂以及水产养殖智慧化的进程。

(2)构建基于开放式养殖池的视频帧数据集,并使用旋转、添加高斯噪声及色彩抖动等增强方式对数据进行增强,提高样本的多样性及随机性。

(3)设计了 CBAM - MV2 模块,在原网络的第 2 层和第 5 层的残差模块中添加 CBAM 注意力模块,对鱼群摄食视频的通道和空间信息进行采样,提高模型对视频的特征能力。同时融合 BiLSTM 双向循环网络,实现了端到端的训练输出,提高了 MobileViT - CBAM - BiLSTM 模型对视频序列的解析能力,鲁棒性以及泛化性。

(4)利用准确率、微 F1 值、宏 F1 值和加权 F1 值 4 个综合指标对模型进行性能评估,改进后的 MobileViT - CBAM - BiLSTM 模型准确率达到 98.61%。

## 参 考 文 献

- [1] 王鲁民,王忠秋.海洋捕捞技术、渔具渔法研究进展与趋势[J].水产学报,2023,47(11):204-214. WANG Lumin, WANG Zhongqiu. Review of marine fishing equipment and technology[J]. Journal of Fisheries of China, 2023, 47(11):204-214. (in Chinese)
- [2] 王炳旭.近海捕捞渔业发展现状问题与对策[J].南方农业,2020,14(23):114-115.
- [3] HARSIJ M, KANANI H G, ADINEH H. Effects of antioxidant supplementation (nano selenium, vitamin C and E) on growth performance, blood biochemistry, immune status and body composition of rainbow trout (*Oncorhynchus mykiss*) under sub-lethal ammonia exposure[J]. Aquaculture, 2020, 521: 734942.
- [4] CHEN L, YANG X, SUN C, 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.
- [5] 刘伟,蒋明,吴凡,等.饲料能量密度和投喂水平对吉富罗非鱼生长和健康的影响[J].水生生物学报,2022,46(1):58-68.

- LIU Wei, JIANG Ming, WU Fan, et al. Effects of dietary energy density and feed ration on growth and health in genetically improved farmed tilapia, *Oreochromis niloticus*[J]. Acta Hydrobiologica Sinica, 2022, 46(1): 58 - 68. (in Chinese)
- [6] ZHOU C, XU D, 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.
- [7] YANG X, ZHANG S, LIU J, et al. Deep learning for smart fish farming: applications, opportunities and challenges[J]. Reviews in Aquaculture, 2021, 13(1): 66 - 90.
- [8] 郑金存, 赵峰, 林勇, 等. 基于近红外深度图的游泳型鱼类摄食强度实时测量[J]. 上海海洋大学学报, 2021, 30(6): 1067 - 1078.
- ZHENG Jincun, ZHAO Feng, LIN Yong, et al. Evaluation of fish feeding intensity in aquaculture based on near-infrared depth image[J]. Journal of Shanghai Ocean University, 2021, 30(6): 1067 - 1078. (in Chinese)
- [9] LU Y C, TUNG C, KUO Y F. Identifying the species of harvested tuna and billfish using deep convolutional neural networks [J]. ICES Journal of Marine Science, 2020, 77(4): 1318 - 1329.
- [10] JALAL A, SALMAN A, MIAN A, et al. Fish detection and species classification in underwater environments using deep learning with temporal information[J]. Ecological Informatics, 2020, 57: 101088.
- [11] LI D, WANG Z, WU S, et al. Automatic recognition methods of fish feeding behavior in aquaculture: a review [J]. Aquaculture, 2020, 528: 735508.
- [12] GUO Q, YANG X, ZHOU C, et al. Fish feeding behavior detection method based on shape and texture features[J]. Journal of Shanghai Ocean University, 2018, 27(2): 181 - 189.
- [13] HU X, LIU Y, ZHAO Z, et al. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO - V4 network[J]. Computers and Electronics in Agriculture, 2021, 185: 106135.
- [14] 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 CSAE, 2020, 36(13): 158 - 164. (in Chinese)
- [15] FENG S, YANG X, LIU Y, et al. Fish feeding intensity quantification using machine vision and a lightweight 3D ResNet - GloRe network[J]. Aquacultural Engineering, 2022, 98: 102244.
- [16] UBINA N, CHENG S C, CHANG C C, et al. Evaluating fish feeding intensity in aquaculture with convolutional neural networks [J]. Aquacultural Engineering, 2021, 94: 102178.
- [17] 唐宸, 徐立鸿, 刘世晶. 基于光流法的鱼群摄食状态细粒度分类算法[J]. 农业工程学报, 2021, 37(9): 238 - 244.
- TANG Chen, XU Lihong, LIU Shijing. Fine-grained classification algorithm of fish feeding state based on optical flow method [J]. Transactions of the CSAE, 2021, 37(9): 238 - 244. (in Chinese)
- [18] 徐立鸿, 黄薪, 刘世晶. 基于改进 LRCN 的鱼群摄食强度分类模型[J]. 农业机械学报, 2022, 53(10): 236 - 241.
- XU Lihong, HUANG Xin, LIU Shijing. Recognition of fish feeding intensity based on improved LRCN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 236 - 241. (in Chinese)
- [19] ZENG Y, YANG X, PAN L, et al. Fish school feeding behavior quantification using acoustic signal and improved Swin Transformer[J]. Computers and Electronics in Agriculture, 2023, 204: 107580.
- [20] MEHTA S, RASTEGARI M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer [J]. arXiv Preprint, arXiv:2110.02178, 2021.
- [21] HUANG Z, XU W, YU K. Bidirectional LSTM - CRF models for sequence tagging[J]. arXiv Preprint, arXiv:1508.01991, 2015.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C] // Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.
- [23] 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, 2018: 4510 - 4520.
- [24] JIANG Y, ZHAO T, CHAI Y, et al. Bidirectional LSTM - CRF models for keyword extraction in Chinese sport news [J]. MIPPR 2019: Pattern Recognition and Computer Vision, 2020: 11430.
- [25] SOKOLOVA M, LAPALME G. A systematic analysis of performance measures for classification tasks [J]. Information Processing & Management, 2009, 45(4): 427 - 437.
- [26] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735 - 1780.
- [27] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization [J]. arXiv Preprint, arXiv:1409.2329, 2014.
- [28] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [J]. arXiv Preprint, arXiv:1412.3555, 2014.