

基于人工智能的鱼类行为识别研究综述

彭秋珺^{1,2} 李蔚然^{1,2} 李振波^{1,2}

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

摘要: 鱼类行为识别对于生态学、水产养殖、渔业资源管理等方面具有重要意义,可以通过其行为模式判断其生长发育状况和活动水平,并间接评估环境因素对其影响,以减少鱼类生长应激反应,提高资源利用效率,为水产养殖的智能化发展奠定基础。近年来,基于人工智能技术的鱼类行为识别方法受到广泛关注,其具有无损性、低成本等优势。本文综述了近5年基于卷积神经网络、循环神经网络、双流卷积神经网络等人工智能方法的鱼类行为识别技术,对鱼类行为识别方法及数据集进行了归纳与分析,在此基础上,对未来的研究进行讨论与展望。

关键词: 鱼类行为识别; 人工智能; 数据集; 水产养殖

中图分类号: S126 **文献标识码:** A **文章编号:** 1000-1298(2023)S1-0283-13

Review of Fish Behavior Recognition Methods Based on Artificial Intelligence

PENG QiuJun^{1,2} LI Weiran^{1,2} LI Zhenbo^{1,2}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. National Innovation Center for Digital Fishery, China Agricultural University, Beijing 100083, China)

Abstract: With the rapid development and expansion of global aquaculture and the diversification of farming models, the scale, intelligence, and informatization of the aquaculture industry have become trends in its development. Fish behavior recognition is of significant importance in ecology, aquaculture, and fisheries resource management. It enables the assessment of fish growth, developmental status, and activity levels based on their behavioral patterns, indirectly evaluating the impact of environmental factors. This can help reduce stress responses in fish growth, improve resource utilization efficiency, and lay the foundation for intelligent development in aquaculture. Traditional fish behavior identification mainly relies on manual observation and recording, which consumes a considerable amount of time and effort and is subject to subjectivity and uncertainty. In recent years, fish behavior recognition methods based on artificial intelligence get extensive attention, is lossless, such as low cost advantage. The fish behavior recognition technologies were reviewed based on artificial intelligence over the past five years, including convolutional neural networks, recurrent neural networks, and two-stream convolutional neural networks. It also provided a summary and analysis of fish behavior recognition methods and datasets. Based on these foundations, an outlook on future research directions was discussed and provided.

Key words: fish behavior recognition; artificial intelligence; datasets; aquaculture

0 引言

近年来,随着全球水产养殖业的迅猛发展及养殖场规模不断扩大,水产养殖产业的规模化、智能化和信息化已成为水产养殖业发展的趋势^[1]。据联合国粮食及农业组织统计,中国已成为渔业和水产

养殖总产量最大的生产国^[2]。2022年中国水产品总产量达 6.866×10^7 t,其中水产养殖产量 5.565×10^7 t,占全国水产品总量的81.1%,养殖面积 7.11×10^6 hm²。水产养殖业能够提供超过人均动物蛋白摄入量的20%,带动农村经济的发展,提高农民收入和生活水平,但也面临着养殖环境污染、养殖成本上

收稿日期: 2023-06-30 修回日期: 2023-08-24

基金项目: 广东省重点领域研发计划项目(2020B0202010009)

作者简介: 彭秋珺(1998—),女,硕士生,主要从事计算机视觉研究,E-mail: pengqiuJun@cau.edu.cn

通信作者: 李振波(1978—),男,教授,博士生导师,主要从事计算机视觉研究,E-mail: lizb@cau.edu.cn

升等诸多挑战^[3]。

鱼类在养殖过程中容易受到生理、心理或环境变化的影响,表现出饥饿、互残、疾病等多种异常行为状态。因此,对鱼类的摄食、浮头等行为的监测与识别有助于实现健康状况的实时监控,可为养殖者提供关键的信息和工具,以优化养殖管理策略、提高生产效率和保障鱼类福利。同时,通过对鱼类行为的观察和分析,可以更好地保护水体和生态环境,实现生态健康养殖模式,促进水产养殖的可持续发展。自 2020 年农业农村部启动实施水产绿色健康养殖技术推广“五大行动”以来,骨干基地基本实现生态健康养殖模式全覆盖、养殖尾水循环综合利用或达标排放全覆盖,水产养殖用兽药使用量同比减少 9%,配合饲料替代率平均达到 84%。此外,对于鱼类行为的观测可以更好地管理渔业资源、保护海洋生态环境,并为海洋科学研究提供有价值的数据和信息,有助于实现可持续的海洋利用和管理,保护海洋生态系统的健康与稳定。2020 年 3 月,谷歌公司的 X 实验室宣布其最新的 Moonshot 项目——Tidal,开发出水下摄像系统,通过人工智能技术跟踪成千上万条鱼类的生长情况,从而保护海洋的良性生态环境,并以可持续的方式造福人类。

传统的鱼类行为识别主要依靠人工观察和记录,但这种方法需要大量的时间和精力,并且存在主观性和不确定性。随着人工智能技术的不断发展,

越来越多的研究人员开始利用人工智能技术对鱼类行为进行识别和分析,以提高研究效率,并且可以更加客观地分析和解释鱼类行为^[4]。这种技术手段通过数字化处理和自动化识别的方式,可以实现对鱼类健康状况的实时监控,提高养殖效率和产量,降低养殖成本,保证产品的质量和安全性,提高水产养殖的智能化和自动化管理水平。

本文将综述现阶段基于人工智能的鱼类行为识别方法以及近 5 年在鱼类行为识别研究中的应用进展,同时总结鱼类行为研究相关的数据集。最后探讨人工智能技术在鱼类行为识别研究中遇到的问题,并对未来的发展方向进行展望。

1 基于人工智能的行为识别方法

行为识别是指利用人工智能技术对鱼类的行为进行自动识别和分类。大多数养殖场的行为监测仍然依靠人工巡视和手动记录,效率低下,容易出现漏检或误判。近年来,以深度学习为代表的新一代人工智能因其强大的学习能力和特征表达能力得到快速发展和应用,被逐渐应用于鱼类行为识别当中。按照人工神经网络结构的不同可以分为基于卷积神经网络(Convolutional neural networks, CNN)、循环神经网络(Recurrent neural networks, RNN)、CNN 与 RNN 结合的网络、双流卷积神经网络(Two-stream convolutional networks)等行为识别方法^[5],各类方法的对比分析如表 1 所示。

表 1 鱼类行为识别方法对比分析
Tab.1 Comparative analysis of fish behavior recognition methods

代表性方法	网络结构	优势	劣势
CNN	由多个卷积层和池化层组成	结构简单,速度较快,可自动学习行为特征	数据需求量大,计算复杂性高,需要较大的计算资源和存储空间
RNN	由一个或多个 RNN 单元组成	可以捕捉鱼类行为的时间依赖关系及上下文信息	计算效率较低,难以捕捉长期时间依赖关系
CNN 与 RNN 结合	由一部分 CNN 用于提取图像特征,然后输入到一个或多个 RNN 层中	建模能力较强,识别性能较高	层数浅,易掉帧
双流卷积神经网络	由一个时间流和一个空间流组成	同时学习图像和时间序列的特征	时间规模和实时性较差

1.1 基于 CNN 的行为识别方法

卷积神经网络(CNN)能够自动学习数据中的特征表示,适用于处理图像和视频数据。通过在 CNN 的网络结构中引入卷积层和池化层,可以有效地提取图像或视频帧的空间特征。这些特征可以捕捉到行为中的纹理、形状和物体位置等语义信息,为后续的分类或回归任务提供依据。CNN 的 3D 卷积结构或时序模型可以对时间序列数据中行为的演化和时间关系进行建模,CNN 可以通过这种方式进行连续的特征提取和预测,从而实现行为识别的任务。

行为识别涉及到对行为的时空关系进行建模和理解,CNN 通过局部感知和参数共享的方式,能够有效地捕捉视频中的时空关系。例如,卷积核的感受野可以覆盖到鱼类图像或视频中的局部区域,从而捕捉到鱼类局部动作的时空关系。

按照卷积的维度不同,CNN 又可以分为 2 维 CNN 和 3 维 CNN。其中,2 维 CNN 主要是针对减少网络参数和提取时间光流信息来改进。DeepVideo^[6]提出在每个视频帧上单独使用一个 2 维的 CNN 模型,并研究了几种时间连接模式,以

学习用于视频行为识别的时空特征。HE 等^[7]提出残差网络,通过不同层间的跳跃连接以减少网络参数,解决退化问题。而 3 维 CNN 可以较好地保留原始通道的信息,卷积结果包含时间维度,在视频处理中应用广泛。TRAN 等^[8]提出的 C3D 采用精简网络的结构和缩小卷积核的方式,具有较强的泛化能力,可作为各种视频任务的通用特征提取器。CARREIRA 等^[9]提出的 I3D 将一个视频片段作为输入,通过堆叠的 3 维卷积层转发,将成熟的图像分类体系用于 3 维 CNN,对于模型权值采用光流网络初始化方法,将 ImageNet 预先训练的 2 维模型权值膨胀到 3 维模型中。QIU 等^[10]提出的 P3D 基于小卷积核运用串行、并行、残差连接方法,通过 ResNet 构建网络。TRAN 等^[11]提出的斜(2+1)维将 2 维空间卷积和 1 维时间卷积关联起来,使得对于时空信息的处理有了非线性化的提升。

在水产养殖方面,2 维 CNN 和 3 维 CNN 都有广泛的应用。ZHOU 等^[12]利用 LeNet5 CNN 方法,对鱼类摄食行为的强度自动分级来评估鱼的食欲,在提高资源利用率、降低饲料成本的同时确保其正常生长,为指导生产实践奠定了理论基础,其模型方法如图 1 所示。HAN 等^[13]提出了一种基于 CNN 和时空信息融合的群体行为识别方法,旨在利用卷积神经网络在图像识别和状态分类方面的突出性能,模拟人脑处理视觉信号时腹侧通道和背侧通道的注意机制,获得了较好的摄食行为识别率。HU 等^[14]提

出了一种改进 YOLOV4 网络的水产养殖未食用饲料颗粒检测模型来对水中残饵进行检测,可以评估鱼群摄食情况及饲料的利用效率,帮助养殖者制定更科学的饲养计划。HU 等^[15]提出了一种基于水下成像系统和深度学习框架的无损、快速、低成本的方法,使用改进的 YOLOV3-Lite 检测混合养殖系统中的鱼类行为,实现对鱼类行为的高精度实时监测。孙龙清等^[16]将 RepVGG^[17]与 GhostNet^[18]相结合,构建了轻量化的基于鱼类行为的水质参数分析模型 G-RepVGG,提出了计算量较少、推理速度快、更适合水质快速监测的 Cheap Ghost 操作以及计算量大、精确率高、更适合水质精确监测的 Expensive Ghost 操作,从而在保证较高精度的前提下依旧具有较高的推理速度,且通过分析在不同温度以及溶解氧下的摄食行为,实现对鱼类所处的温度以及溶解氧浓度的水质监测。HU 等^[19]利用 C3D 深度网络模型对复杂环境下采集的鱼类行为数据集进行异常行为识别,实验取得了较好的识别效果,可用于实际水产养殖中视频监控鱼类的异常行为。UBINA 等^[20]引入了 I3D 模型,将原始 RGB 视频转换为光流图像序列,再将其输入到 3 维 CNN 中,以评估或分类鱼类摄食强度水平。基于 CNN 的鱼类行为识别方法如表 2 所示。

CNN 结构简单且具有较快的运行速度,可以通过多个卷积层和池化层从原始图像数据中自动学习鱼类的形状、纹理和姿态等的特征,有助于更准确地

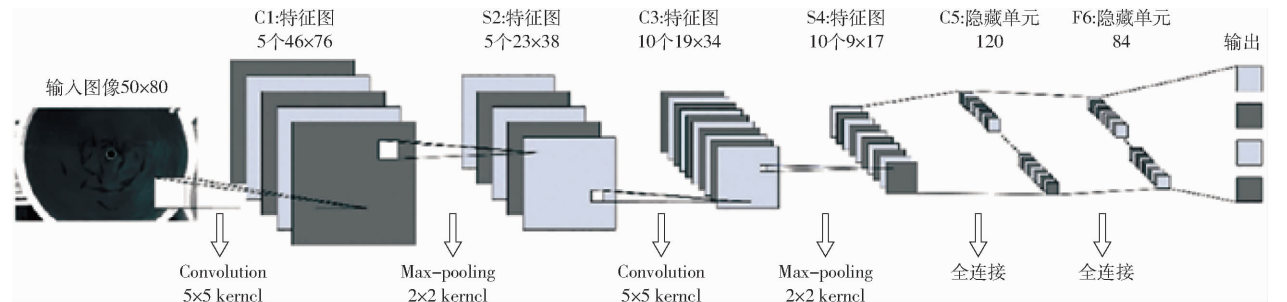


图 1 基于 CNN 的鱼类行为识别方法示例

Fig. 1 Example of fish behavior recognition method based on CNN

表 2 基于 CNN 的鱼类行为识别方法

Tab. 2 Fish behavior recognition methods based on CNN

代表性成果	方法	研究对象	行为类型	准确率/%
ZHOU 等 ^[12]	LeNet5 CNN	罗非鱼	摄食行为	90
HAN 等 ^[13]	CNN 和时空信息融合	斑马鱼	摄食行为	82.5
HU 等 ^[14]	改进的 YOLO-V4	鲑鱼	摄食行为	92.61
HU 等 ^[15]	改进的 YOLOV3-Lite	鲫鱼、黄鲇	摄食、缺氧行为	89.7
孙龙清等 ^[16]	G-RepVGG	斑石雕	不同温度及溶解氧下的摄食行为	96.21 (Cheap Ghost) 97.63 (Expensive Ghost)
HU 等 ^[19]	C3D	大型鱼群	水质变化时聚集和浮头的异常行为	92-97
UBINA 等 ^[20]	I3D	海鲈、军曹鱼	摄食行为	95

识别复杂的鱼类行为。但 CNN 通常需要大量的标注数据来进行训练,收集和标注数据的成本较高,且模型往往具有较大的参数量和计算复杂性,需要较高的计算资源和存储空间。

1.2 基于 RNN 的行为识别方法

循环神经网络(RNN)是一种用于处理序列数据的神经网络模型,其网络结构由一个或多个 RNN 单元组成,通过在时间上的循环连接,使得网络可以处理时序数据的依赖关系。RNN 的核心思想是引入循环结构,具有反馈连接,可以处理较长的序列输入,允许信息在网络中传递。在 RNN 中,每个时间步都有一个隐藏状态,它在处理每个输入时更新,并将先前时间步的信息传递给当前时间步。这种隐藏状态的反馈机制使 RNN 能够对序列中的上下文和时间依赖关系进行建模,从而在许多序列建模任务中表现出色。

HOCHREITER 等^[21]提出长短期记忆(Long short term memory,LSTM)单元,主要包括忘记阶段、选择记忆阶段和输出阶段 3 部分。传统的 RNN 更新模块有一层 tanh 层作为激活层,而 LSTM 有 4 个不同的层相互作用,在长序列的情况下表现更好。余兴^[22]利用快进连接的时空融合模型为基础构建

深层 LSTM。LI 等^[23]利用运动注意力机制的 Video-LSTM,并引入 CNN 来提取视频帧中的空间相关性。因此,RNN 适用于处理时间序列信息明显的视频图像,但由于 LSTM 的结构约束,在 CNN 复杂时建模的准确度会随之降低。

随着 RNN 的迅速发展,鱼类行为识别的研究也有了一定的进展。ZHAO 等^[24]开发了一种基于改进运动影响图和 RNN 的集约化养殖鱼群突发聚集和逃散的局部异常行为检测、定位和识别方法,用于评估水生环境中微生物入侵和食物匮乏引起的鱼类行为异常状况,可以让养殖人员及时发现异常从而采取措施,该方法如图 2 所示。刘世晶等^[25]提出一种基于帧间光流特征和改进 RNN 的草鱼摄食状态分类方法,利用改进的 RNN 网络实现摄食状态分类,增强投喂状态识别的时间相关性,为鱼类精准投喂技术研究提供参考。WEI 等^[26]利用改进的动能模型(Modified kinetic energy model,MKEM)和定制的 RNN,提出了一种基于鱼类行为时空特征的新型实用性鱼类食欲评估方法,利用 MKEM 提取并量化空间行为特征,使用定制的 RNN 学习鱼类摄食行为的时间特征,并对鱼类食欲水平进行分类。基于 RNN 的鱼类行为识别方法如表 3 所示。

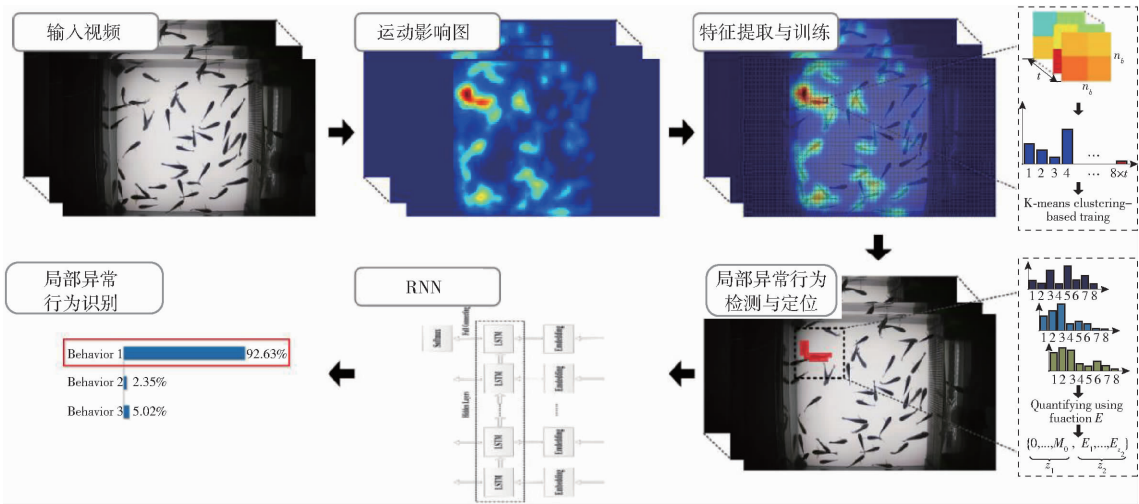


图 2 基于 RNN 的鱼类行为识别方法示例

Fig.2 Example of fish behavior recognition method based on RNN

表 3 基于 RNN 的鱼类行为识别方法				
Tab.3 Fish behavior recognition methods based on RNN				
代表性成果	方法	研究对象	行为类型	准确率/%
ZHAO 等 ^[24]	改进的运动影响图 和 RNN	罗非鱼	突发聚集和	89.89
			逃散的局部 异常行为	
刘世晶等 ^[25]	帧间光流特征和 改进 RNN	草鱼	摄食行为	91.00
WEI 等 ^[26]	MKEM、RNN	鲈鱼	摄食行为	98.31

RNN 在处理鱼类行为的时间特征方面有不错

的效果,可以捕捉鱼类行为的时间依赖关系,从而更好地理解 and 识别鱼类的运动轨迹和动作。同时,RNN 也可以捕捉鱼类行为中的上下文信息,从而更好地理解当前的鱼类行为,提升水产养殖的生产效益。但目前 RNN 在处理长时间序列时会产生大量的计算从而导致计算效率降低,且可能会出现梯度消失或梯度爆炸等问题,导致模型难以捕捉长期时间依赖关系。

1.3 基于 CNN 与 RNN 结合的行为识别方法

由于 CNN 和 RNN 具有不同的特性,将两者结

合可以同时捕捉空间和时间信息。其网络结构通常是由一部分 CNN 用于提取图像特征,然后将提取的特征输入到一个或多个 RNN 层中用于处理时序信息,具有较高的行为识别性能。

DONAHUE 等^[27]将传统 CNN 网络和 LSTM 相结合,提出了长期循环卷积网络(Long-term recurrent convolutional networks, LRCN),能够同时对视频序列或单帧图像进行预测。YUE-HEI 等^[28]提出了 2 种能够将帧级 CNN 输出聚合成视频级预测的视频分类方法。一种方法探索了各种卷积时间特征池架构,另一种方法将视频建模为有序的帧序列,采用 CNN 输出后连接到 LSTM 单元的 RNN。SRIVASTAVA 等^[29]使用 LSTM 网络来学习视频序列的表示,以执行不同的任务。WU 等^[30]提出了一种用于视频分类的混合深度学习框架,该框架能够对视频中的静态空间信息、短期运动以及长期时间线索进行建模。

同时引入 CNN 和 RNN 也有助于提升鱼类行为识别模型的精度,更好地捕捉时空信息。其中,现有的大部分模型是将 CNN 与 LSTM 结合来进行鱼类的行为识别,模型方法如图 3 所示。徐立鸿等^[31]提出了一种基于改进 LRCN 的鱼类摄食强度分类模型,将注意力机制 SE(Squeeze-and-excitation)^[32]模块嵌入卷积神经网络中,通过 SE-CNN 网络提取视频帧的特征,输入至双层门控递归单元(Gate recurrent unit, GRU)^[33]网络中,最后通过全连接分类层得出视频类别,为自动精准投饵提供参考。DU 等^[34]采用了 CNN 结合 LSTM 的方法,提出了一种基于 ResNet50-LSTM 的方法来识别 5 种鱼类繁殖行为,将 ResNet50 通过多重卷积提取到的不同维度的输入数据特征输入到 LSTM 分类模型中,从而能够提取时间序列的行为信息,为智慧水产养殖提供了新的策略。MÅLØY 等^[35]提出了一种双流循环网络(Dual-stream recurrent network, DSRN),通过使用空间网络、3D 卷积运动网络和 LSTM 循环分类网络来组合空间和运动时间信息,自动捕获鲑鱼在游泳

过程中的时空行为。基于 CNN 与 RNN 结合的鱼类行为识别方法如表 4 所示。

表 4 基于 CNN 与 RNN 结合的鱼类行为识别方法

Tab. 4 Fish behavior recognition methods based on CNN and RNN

代表性成果	方法	研究对象	行为类型	准确率/%
徐立鸿等 ^[31]	改进的 LRCN	草鱼、鲫鱼	摄食行为	97.00
DU 等 ^[34]	CNN、LSTM	斑石雕	追逐、产卵、聚集、摄食、正常游泳时的行为	98.52
MÅLØY 等 ^[35]	DSRN	鲑鱼	摄食行为	80.00

将 CNN 和 RNN 两者结合的鱼类行为识别方法在 CNN 中进行图像特征提取的同时,利用 RNN 的序列建模能力学习时序信息。该方法具有较强的建模能力,能够更全面地学习到鱼类行为的时空特征,从而提高识别性能,但仍存在层数浅、易掉帧的问题。

1.4 基于双流卷积神经网络的行为识别方法

光流可以有效地去除没有变化的背景,与使用原始的 RGB 图像作为输入相比可以简化学习问题,因此 SIMONYAN 等^[36]提出了双流网络,包括一个空间流和一个时间流,缩小了人工智能方法和传统手工制作特性之间的差距。其中空间信息和时间信息的处理流程可以分别进行,最后将 2 个流的输出特征进行融合或连接,用于行为识别任务。

双流网络使用相对较浅的网络架构,因此许多研究人员将双流网络扩展到了更深层次。WANG 等^[37]引入跨模态初始化、同步批处理规范化、角落种植和多尺度种植数据增强等方法,防止深层网络过拟合。而双流网络中存在 2 个流,需要将 2 个网络的结果进行合并以获得最终的预测结果。FEICHTENHOFER 等^[38]通过引入 2 个流之间的剩余连接,将 ResNet 推广到时空域。由于视频本质上是一个时间序列,研究人员探索了 RNN 特别是 LSTM 在视频中的时间建模。然而,光流网络仍然不能捕捉到远距离的时间信息。因此,WANG 等^[39]提出了一种时域分段网络(Temporal segment network, TSN)来进行视频级行为识别。TSN 能够对整个视频进行长距离的时间结构建模。此外,这种稀疏采样策略降低了长视频序列的训练成本,但是却保留了相关信息。为了提高 TSN 的时间推理能力,提出了时间关系网络(Temporal relation network, TRN)^[40],用于在多个时间尺度上学习和推理视频帧之间的时间相关性。FEICHTENHOFER 等^[41]提出 SlowFast 网络,以低帧速率运行的慢速路径来捕获空间语义,以高帧速率运行的快速路径来捕获时

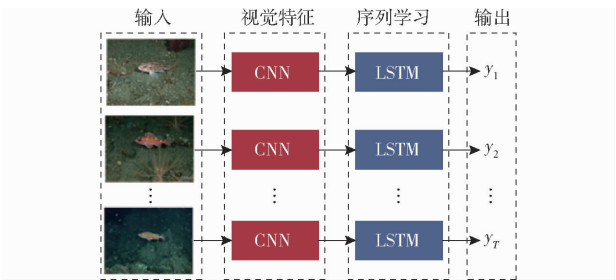


图 3 基于 CNN 与 RNN 结合的鱼类行为识别方法示例

Fig. 3 Example of fish behavior recognition method based on CNN and RNN

间分辨率。

在近两年的研究当中,基于双流卷积神经网络的鱼类行为识别方法也较常见,结构如图 4 所示。WANG 等^[42]通过改进双流卷积神经网络,以 RGB 和光流视频片段作为输入,提出了基于深度学习的具有时空和运动信息的双流三维卷积神经网络 (Dual-stream 3D convolutional neural network,

DSC3D),为实时从视频序列中自动识别鱼类行为提供了一种有效的策略。DITRIA 等^[43]使用密集的光流来评估水下镜头中的像素运动,使用时空滤波来提高准确性,并使用深度学习算法对放牧行为进行分类,为当前耗时的手动数据提取方法提供了一种替代方案。基于双流卷积神经网络的鱼类行为识别方法如表 5 所示。

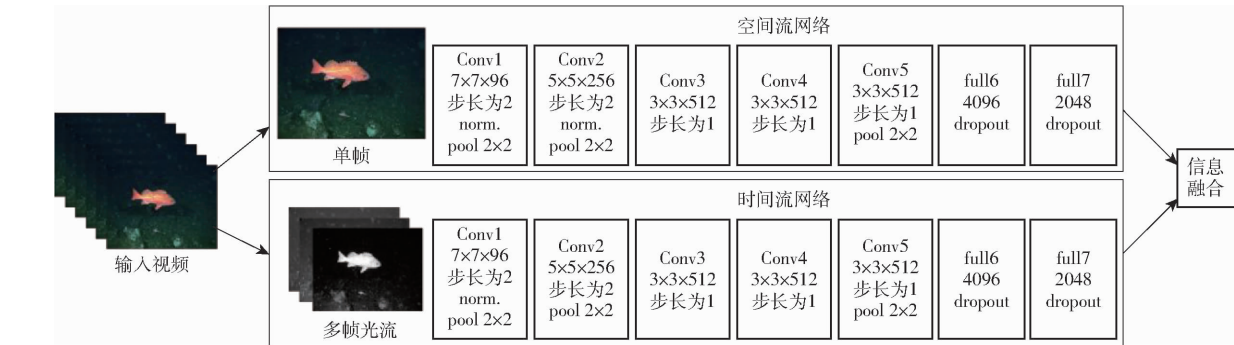


图 4 基于双流卷积神经网络的鱼类行为识别方法示例

Fig. 4 Example of fish behavior recognition method based on two-stream convolutional network structure

表 5 基于双流卷积神经网络的鱼类行为识别方法				
Tab.5 Fish behavior recognition methods based on two-stream convolutional networks				
代表性成果	方法	研究对象	行为类型	准确率/%
WANG 等 ^[42]	DSC3D	斑石鲷	进食、缺氧、惊	95.79
			吓、体温过低时	
DITRIA 等 ^[43]	光流法	三尖鲷	游泳行为	92.00

通过将空间流和时序流的分别处理,双流卷积神经网络能够充分利用图像和时序数据的信息,更全面地理解鱼类行为。该方法精确度高,扩展性强,对复杂的行为识别任务有一定优势,但与其他模型相比时间规模和实时性较差,因此适合在对实时性要求不高的情况下使用。

1.5 其他行为识别方法

双流网络通过外观和运动信息进行行为识别,除此之外利用其他因素也能达到很好的识别效果。例如在没有场景背景的情况下,可以通过观察姿势来识别行为,P-CNN^[44]提出沿身体部位轨迹聚合运动和外观信息,是最早成功利用姿态提高行为识别效果的深度学习方法之一。物体信息也是一个重要的因素,WU 等^[45]提出利用物体特征和场景特征来进行行为识别,WANG 等^[46]更进一步使用端到端网络进行训练。音频信号通常作为视觉信息的补充与视频信号一起出现,WU 等^[47]引入了一种多流框架,该框架整合了视频中的空间、短期运动、长期时间和音频。在 RGB-D 视频行为识别领域,Depth2Action^[48]使用现有的深度估计器从视频中提取深度信息,并将其用于行

为识别。

在鱼类的行为识别当中,大多数方法是利用目标跟踪形成鱼类运动轨迹,通过对轨迹的分析来研究鱼类的具体行为。WANG 等^[49]提出了一种结合关键点检测、定向循环图 (Directed cycle graph, DCG)^[50]、鱼类跟踪和动态时间规整 (Dynamic time warping, DTW)^[51]的水下鱼类异常行为检测方法,对鱼的 8 个身体部位进行预先标记,有助于提前探测水下鱼类生物异常行为,从而规划和实施早期对策。CHENG 等^[52]提出了一种结合 3 维运动轨迹合成和集成学习的生物水质监测方法,使用 MobileNet_SSD 来检测鱼类并标记鱼的具体位置,采用核化相关滤波器 (Kernelized correlation filters, KCF) 跟踪算法对运动和遮挡过程中丢失或被遮挡的鱼体运动目标进行补偿,能够在鱼类 3 维轨迹像素坐标下有效、准确地反映水质。罗冰^[53]利用机器学习和神经网络的方法,提出了基于 PSO_BPNN 的鲫鱼浮头数量预测模型,实现对养殖环境中鲫鱼浮头行为的实时监测和预警。张笑华等^[54]发展基于 Pareto 多目标人工鱼群算法 (Multi-objective artificial fish swarm algorithm, MO-AFSA),改进人工鱼群算法的追尾和觅食行为,解决结构健康监测中传感器位置多目标优化的问题。COOK 等^[55]验证了自适应分辨率成像声呐 (Adaptive resolution imaging sonar, ARIS) 在测量 4 种具有不同游泳方式的鱼类的效用,验证了成像声呐在渔业研究不同领域当中的有效性。罗毅等^[56]提出一种基于 One-Class SVM 模型的青鲈鱼异常行为识别方法,将采集到的暴露在不同类型和

浓度特征污染物下的青鲮鱼行为强度信号作为经验数据,利用直方图统计和主成分分析对行为强度数据进行降维,实现行为特征提取。该方法可对水源突发性水质污染进行更有效的监测和预警,也可为水污染应急决策提供技术支撑。其他鱼类行为识别方法如表 6 所示。

表 6 其他鱼类行为识别方法
Tab.6 Other fish behavior recognition methods

代表性成果	方法	研究对象	行为类型	准确率/%
WANG 等 ^[49]	关键点检测、定向循环图	罗非鱼	剧烈环境变化下的异常行为	92.80 (Faster RCNN)
	(DCG)、鱼类跟踪和动态时间规			88.00 (YOLO v4)
	整 (DTW)			81.80 (EfficientDet-D1)
CHENG 等 ^[52]	KCF 跟踪算法	红鲫鱼	游泳行为	95.00
罗冰 ^[53]	PSO_BPNN	鲫鱼	浮头行为	97.14 (测试集)
张笑华等 ^[54]	Pareto 多目标人工鱼群算法	人工鱼群	追尾、觅食行为	85.04 (验证集)
COOK 等 ^[55]	ARIS	蓝鳕鱼、黄眼鲷鱼、鲷鱼和鲈鱼	游泳行为	
罗毅等 ^[56]	One-Class SVM	青鲮鱼	在不同类型和浓度特征污染物下的异常行为	有机污染物:10 min 内完成预警 重金属污染物:1 h 内完成预警

2 鱼类行为识别数据集

数据集是基于数据驱动的人工智能方法的关键,鱼类行为识别数据集是科学研究鱼类行为及技术应用的基础,通过在鱼类数据集上进行测试,研究人员可以客观地比较不同算法的性能和准确性。数据集按照采集方式的不同分为公开数据集和专用数据集两种。

2.1 鱼类行为识别公开数据集

鱼类行为识别的公开数据集可以用于评估不同鱼类行为识别方法的性能,任何研究人员都可以使用相同的鱼类数据集进行实验,从而验证和复现其他研究的结果。通过研究人员的合作和数据共享,

有助于推动该领域的研究进展,在避免重复收集数据工作的同时提高研究效率。

2010 年 10 月 1 日至 2013 年 9 月 30 日期间,研究人员从中国台湾南湾海峡、兰屿岛和后壁湖的水下观景台收集鱼类图像形成数据集 Fish4Knowledge^[57],收集到的数据可以用于鱼类的检测和识别、目标跟踪及行为分析。该数据集共使用 10 台摄像机,在日光环境中收集到超过 12 h 的水下数据。图像数据集共有 27 370 幅,包含 23 种鱼类,其中不同类型的图像数量有所差别。数据集规模较大,是最典型的鱼类目标公共数据集之一,但数据类型明显不平衡。Fish4Knowledge 数据集部分示例如图 5 所示^[57]。



图 5 Fish4Knowledge 数据集示例
Fig.5 Fish4Knowledge dataset example

LifeCLEF 数据集包含 FishCLEF 2014^[58]数据集、FishCLEF 2015^[59]数据集、SeaCLEF 2016^[60]数据集。其中 FishCLEF 2014 数据集的数据来源是 Fish4Knowledge 水下数据集,数据集中的训练集共有 285 个视频,19 868 个鱼类注释,包含 10 种鱼的类别。FishCLEF 2015 数据集的训练集共有 20 个手动标注的视频,共 15 种鱼的类别,包含 9 162 个标注,超过 20 000 幅图像。与 Fish4Knowledge 数据集相比,FishCLEF 2014 和 FishCLEF 2015 进行了广泛的扩展,考虑了一般的海洋生物,使得训练效果更

好。SeaCLEF 2016 数据集基于 FishCLEF 2014 和 FishCLEF 2015 数据集增加了鲸鱼、海豚、海床、珊瑚等目标,训练集共有 20 个低分辨率视频和 20 000 多幅样本图像。JÄGER 等^[61]对 SeaCLEF2016 数据集进行了标注。

Labeled Fishes in the Wild^[62]数据集由美国国家海洋渔业管理局 (National Oceanic and Atmospheric Administration, NOAA) 提供,通过西南水产科学中心使用部署在远程操作车辆 (Remotely operated vehicles, ROV) 上的前向斜数码相机拍摄,主要包含

鱼类、无脊椎鱼类、海床图像以及相应的标注文件,文件标注出鱼类目标的位置。该数据集可以分为3部分:用于训练和验证鱼类数据的正样本、用于训练和验证非鱼类数据的负样本和测试图像集。其中,正样本当中共有 929 个图像文件,包含 1 005 条具有标记位置和边框的鱼类。负样本共有 3 167 幅图像,包含 147 幅可直接下载的海底图像和 3 020 幅



图 6 Labeled Fishes in the Wild 数据集示例
Fig.6 Labeled Fishes in the Wild dataset example

NCFM (Nature conservancy fisheries monitoring) 数据集来自 Kaggle 主办的“大自然保护渔业监测”全球竞赛。该数据集共有 16 915 幅图像,其中包含 3 777 幅用于训练的图像和 13 138 幅用于测试的图像。由于所有样品图像都是通过安装在不同渔船上的相机从各个角度捕获,NCFM 数据集的场景更加复杂,有四处走动的工人以及地面上的集装箱和渔具。同时,大多数鱼类离相机很远,出现在船底、渔网、人或海水中,并且存在遮挡的现象。由于图像是在不同日期和天气的任何时间拍摄的,在不同图像中光照度的差别很大,且部分图像因为雨水或溅起的海水覆盖相机而导致图像质量低下^[64]。

2018 年和 2019 年“用于环境监测的海洋视频

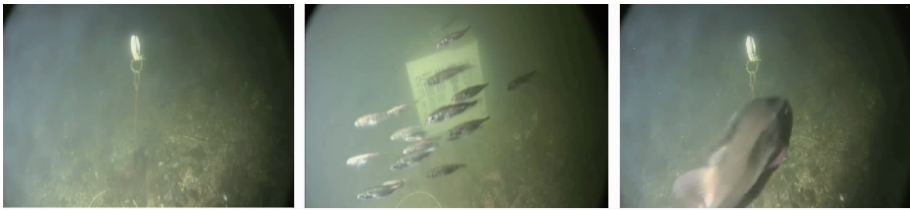


图 7 Brackish 数据集示例
Fig.7 Brackish dataset example

360 像素)的声呐视频序列,并将视频划分为图像。声呐图像中存在严重的噪点和不稳定的视觉特征,因此这些数据非常具有挑战性。此外,鱼类在图像中用单独的点注释标记。该数据集共有来自 30 个视频序列的 537 幅图像,其中包含 24 个视频序列中用于训练和验证的 417 幅图像,以及 6 个视频序列中用于测试的 120 幅图像。部分鱼类行为识别的公开数据集如表 7 所示。

2.2 鱼类行为识别的专用数据集

由于现阶段可用于鱼类行为分析的公开数据集较少,研究人员或相关机构通常自行采集和整理形成用于鱼类研究和分析的专用数据集,通常包含鱼

可从 OpenCV 或目录中获得的图像。测试图像集包含使用 ROV 的高清摄像机捕捉的图像序列,在标注的框架中总共标记了 2 061 条鱼对象。该数据集拍摄质量好,且具有相应的标注,可以用于进一步检测复杂环境中的鱼类以及对不同种鱼类的识别和分类^[63]。Labeled Fishes in the Wild 数据集部分示例如图 6 所示^[62]。

自动化分析”研讨会中发布了 MOUSS (Modular optical underwater survey system) 数据集和 MBARI (Monterey bay aquarium research institute) 水下数据集,这些数据都包含目标的注释,注释用 Bounding box 或者关键点标出。MOUSS 数据集共有 159 幅包含注释的鱼类图像,MBARI 水下数据集共有 666 幅鱼类图像^[65]。

Brackish^[66]数据集于丹麦北部的海峡中拍摄,包含鱼类、螃蟹和其他海洋生物。目标的位置用边界框进行注释,共有 14 518 幅图像,包含 6 个类别的 28 518 个标注信息。Brackish 数据集部分示例如图 7 所示^[66]。

LIU 等^[67]从网上收集了低分辨率(约 360 像素 ×

类的各种信息和属性,如物种、尺寸、重量、年龄、生态环境等。专用数据集可用于研究鱼类的行为,为渔业管理、生态保护和环境决策等提供了重要的参考和依据。图 8 展示了 4 种不同场景下采集的鱼类数据集部分示例^[68]。

专用数据集的优点在于能够提供更具有代表性和真实性的鱼类数据,并且还可以根据具体研究需求进行定制,包括目标物种、观测地点、时间范围等,以满足特定研究目标的需求。然而,专用数据集在收集过程中涉及大量的时间和劳动,需要专业的设备和技术支持,存在一定的挑战和限制。同时,鱼类在自然环境中的行为和特征具有多样性,数据集的样

表 7 鱼类行为识别的公开数据集

数据集	年份	数据类型	类别数量	数据量	下载网址
Fish4Knowledge ^[57]	2014	图像	23	27 370	https://groups.inf.ed.ac.uk/f4k/
FishCLEF 2014 ^[58]	2014	视频	10	285	https://www.imageclef.org/2014/lifeclef/fish
FishCLEF 2015 ^[59]	2015	视频/图像	15	视频:20 图像:20 000 +	http://www.imageclef.org/lifeclef/2015/fish
Labeled Fishes in the Wild ^[62]	2015	视频/图像	7	视频:1 图像:4 096	https://swfscdata.nmfs.noaa.gov/labeled-fishes-in-the-wild
SeaCLEF 2016 ^[60]	2016	视频/图像	15	视频:20 图像:20 000 +	https://www.imageclef.org/lifeclef/2016/sea
NCFM	2017	图像	8	16 915	https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring
MOUSS	2019	图像		159	https://www.viametoolkit.org
MBARI	2019	图像		666	https://www.mbari.org
Brackish ^[66]	2020	图像	6	14 518	https://www.kaggle.com/aalborguniversity/brackish-dataset
Sonar image counting ^[67]	2021	视频/图像		视频:30 图像:537	http://www.soundmetrics.com/

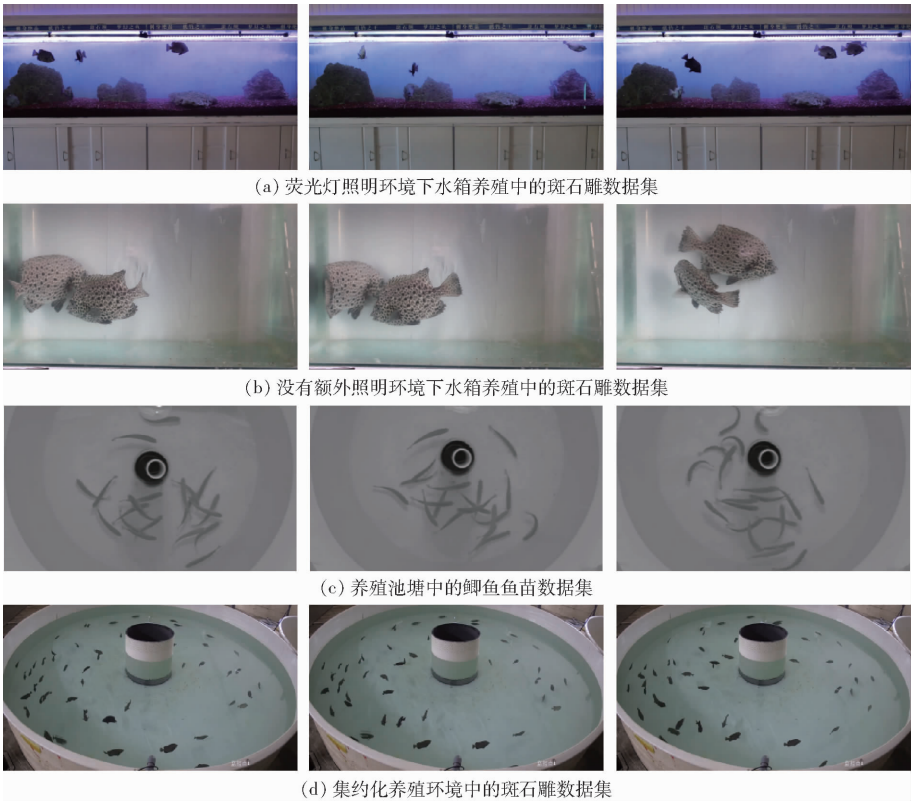


图 8 4 种不同场景下采集的鱼类数据集部分示例

Fig. 8 Four different scenarios of fish datasets example

本量和多样性往往会受到限制。

3 讨论与展望

现阶段鱼类行为识别方法在高精度、低成本上已经取得了一定的成果,但仍然存在一些问题需要解决。本文针对公开数据集构建、多模态信息融合、人工智能模型改进以及养殖模式应用 4 方面作出讨论与展望。

3.1 公开数据集构建

目前可用于鱼类行为识别的公开数据集总量相对较少且缺乏数据标注,数据量较少的数据集图像范围在 150 ~ 700 幅,而数据量较大的数据集图像达 20 000 幅以上。为了推动基于人工智能的鱼类行为识别领域的进一步发展和应用,需要不断收集和整理鱼类行为数据,并建立大规模公开数据集并由专业人士进行数据标注供研究人员使用。其次,鱼类

行为识别需要高质量的数据来训练和验证模型,而当前可用的数据缺乏多样性。现有的鱼类数据集大多在海洋、湖泊等自然环境中采集,涵盖鱼类正常活动时游泳及摄食等行为,且不同类别的样本量并不均衡。未来需要考虑包含不同种鱼类、不同环境条件下的不同行为样本,以及不同时间地点的鱼类个体和群体数据,以尽可能反映真实场景中的多样性,以提高模型的泛化能力和应用范围。同时,大多数公开数据集的数据以相机拍摄的图像为主,视频、声呐图像等其他形式的数据量过少,在未来研究中可对数据类型进行拓展。此外,现有的公开数据集如Brackish数据集中存在数据模糊的情况。随着水下摄像设备的发展,构建高质量、大规模且具有数据标注和数据多样性的鱼类行为公开数据集也将是未来发展的一大趋势。

3.2 多模态信息融合

鱼类行为识别涉及多种感知模态,如视觉、声音、电信号等,融合多种模态数据可以获得更为丰富和全面的行为特征。然而,现阶段的研究大部分是在视觉的基础上开展的,如何将多模态数据融合起来,以提高行为识别的准确性和鲁棒性是一个挑战。由于人工智能技术的迅速发展为多模态融合在鱼类行为识别中的应用提供了广阔的前景,未来可以使用如多模态CNN和多模态RNN的方法,将不同模态的特征进行学习和综合分析,在推动鱼类行为识别的研究和应用的同时也促进鱼类保护和生态环境可持续发展。

3.3 人工智能模型改进

人工智能算法虽然在鱼类行为识别中表现出色,已经达到了较高的准确率,但在面对复杂环境、遮挡、光照变化等因素时会出现不稳定的情况,仍然有进步空间。改进模型的鲁棒性和泛化能力,探索更复杂和高效的人工智能模型,如注意力机制、Transformer等,使其能够在不同环境条件下准确识别鱼类行为,进一步提高鱼类行为识别的性能。其次,部分现有模型的大小限制了模型的特征提取能力。如C3D只有11层网络,但是模型内存占用量有321 MB,相比之下ResNet152的模型内存占用量仅有235 MB。因此,可以在原模型的基础上设计轻量级的网络结构,减少模型的层数和参数量,以达到平衡模型大小和性能的最佳效果。同时,预训练的大模型也是未来研究的一个重要方向,目前一些预训练的大模型如谷歌发布的单模态语言大模型

BERT、百度提出的ERNIE大模型等已经得到了广泛应用。随着数据集扩大和模型优化,用于鱼类行为识别的预训练大模型也将在未来的研究中不断被探索,并推动人工智能技术的进一步发展。此外,随着技术的不断进步和应用的不断拓展,鱼类行为识别的研究可应用于多种领域。未来可针对不同的领域展开研究,如水产养殖、渔业资源管理、水生生态系统监测等,为生态保护和资源管理提供有力的支持和帮助。

3.4 养殖模式应用

随着水产养殖的发展,养殖模式呈现出多样化的趋势,包括工厂化循环水养殖、陆基圆池循环水养殖、室内鱼菜共生养殖、近海新型网箱养殖、深远海设施养殖等。不同养殖模式的环境和条件各异,养殖池大小及养殖规模有所区别,如何提高鱼类行为识别方法的鲁棒性和泛化性,将其应用于不同养殖模式是一个具有挑战性的任务。未来需要综合考虑养殖环境、数据采集方式、鱼类行为特征等因素,根据实际应用场景的不同探究更具有针对性的识别方法,以适应不同养殖模式的需求,实现对鱼类行为的准确识别和智能化管理。现阶段的研究大多是以鲮鱼、罗非鱼、鲫鱼、鲈鱼、草鱼等为研究对象,其他类别的鱼类研究较少。未来可扩大研究对象的种类,并对同一养殖模式下的不同种鱼类或不同养殖模式下的同种鱼类的行为展开更为深入的研究。同时也可以调整养殖密度或鱼类大小,为不同养殖模式提供多元化的研究方法和数据,促进水产养殖的可持续发展。

4 结束语

对现阶段基于人工智能的行为识别方法和技术以及近5年在鱼类行为识别研究中的应用进展按照网络结构的不同分为5类来进行论述,同时总结鱼类行为识别所需要的相关数据集,最后探讨人工智能技术在鱼类行为识别研究中遇到的问题,并对未来的发展方向进行展望。从目前的成果来看,基于人工智能的鱼类行为识别研究在数据集的质量和规模、多种感知模态的融合、预训练的大模型、不同养殖模式下的方法应用等方面仍然还有很大的发展空间。通过进一步创新和改进鱼类行为识别技术,养殖模式将不断丰富和完善,实现高效率的水产绿色健康养殖,保护海洋生态系统的健康与稳定。

参 考 文 献

[1] YANG L, LIU Y, YU H, et al. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior

- analysis: a review[J]. Archives of Computational Methods in Engineering, 2021, 28: 2785 – 2816.
- [2] FAO. The state of world fisheries and aquaculture 2022: towards blue transformation[R]. The State of World Fisheries and Aquaculture (SOFIA), 2022.
- [3] CUI Z, WU J F, YU H, et al. A review of the application of computer vision technology in aquaculture[J]. Marine Science Bulletin, 2018, 20(1): 53 – 66.
- [4] 段延娥, 李道亮, 李振波, 等. 基于计算机视觉的水产动物视觉特征测量研究综述[J]. 农业工程学报, 2015, 31(15): 1 – 11.
- DUAN Yane, LI Daoliang, LI Zhenbo, et al. Review on visual attributes measurement research of aquatic animals based on computer vision[J]. Transactions of the CSAE, 2015, 31(15): 1 – 11. (in Chinese)
- [5] 袁首, 乔勇军, 苏航, 等. 基于深度学习的行为识别方法综述[J]. 微电子学与计算机, 2022, 39(8): 1 – 10.
- YUAN Shou, QIAO Yongjun, SU Hang, et al. A review of behavior recognition methods based on deep learning[J]. Microelectronics & Computer, 2022, 39(8): 1 – 10. (in Chinese)
- [6] KARPATY A, TODERICI G, SHETTY S, et al. Large-scale video classification with convolutional neural networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1725 – 1732.
- [7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 – 778.
- [8] TRAN D, BOURDEV L, FERGUS R, et al. Learning spatiotemporal features with 3D convolutional networks[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 4489 – 4497.
- [9] CARREIRA J, ZISSERMAN A. Quo vadis, action recognition? A new model and the kinetics dataset[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6299 – 6308.
- [10] QIU Z, YAO T, MEI T. Learning spatio-temporal representation with pseudo-3D residual networks[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 5533 – 5541.
- [11] TRAN D, WANG H, TORRESANI L, et al. A closer look at spatiotemporal convolutions for action recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6450 – 6459.
- [12] 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.
- [13] HAN F, ZHU J, LIU B, et al. Fish shoals behavior detection based on convolutional neural network and spatiotemporal information[J]. IEEE Access, 2020, 8: 126907 – 126926.
- [14] 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.
- [15] HU J, ZHAO D, ZHANG Y, et al. Real-time nondestructive fish behavior detecting in mixed polyculture system using deep-learning and low-cost devices[J]. Expert Systems with Applications, 2021, 178: 115051.
- [16] 孙龙清, 王泊宁, 王嘉煜, 等. 基于 G – RepVGG 和鱼类运动行为的水质监测方法[J]. 农业机械学报, 2022, 53(增刊 2): 210 – 218.
- SUN Longqing, WANG Boning, WANG Jiayu, et al. Water quality monitoring based on fish movement behavior and G – RepVGG[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 2): 210 – 218. (in Chinese)
- [17] DING X, ZHANG X, MA N, et al. Repvgg: making vgg-style convnets great again[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13733 – 13742.
- [18] HAN K, WANG Y, XU C, et al. GhostNets on heterogeneous devices via cheap operations[J]. International Journal of Computer Vision, 2022, 130(4): 1050 – 1069.
- [19] HU Z, LI X, XIE X Y, et al. Abnormal behavior recognition of underwater fish body based on C3D model[C] // Proceedings of the 2022 6th International Conference on Machine Learning and Soft Computing, 2022: 92 – 97.
- [20] 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.
- [21] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735 – 1780.
- [22] 余兴. 基于深度学习的视频行为识别技术研究[D]. 成都: 电子科技大学, 2018.
- YU Xing. Video action recognition technology research based on deep learning[D]. Chengdu: University of Electronic Science and Technology of China, 2018. (in Chinese)
- [23] LI Z, GAVRILYUK K, GAVVES E, et al. Videolstm convolves, attends and flows for action recognition[J]. Computer Vision and Image Understanding, 2018, 166: 41 – 50.
- [24] ZHAO J, BAO W, ZHANG F, et al. Modified motion influence map and recurrent neural network-based monitoring of the local unusual behaviors for fish school in intensive aquaculture[J]. Aquaculture, 2018, 493: 165 – 175.
- [25] 刘世晶, 涂雪滢, 钱程, 等. 基于帧间光流特征和改进 RNN 的草鱼摄食状态分类[J]. 水生生物学报, 2022, 46(6): 914 – 921.
- LIU Shijing, TU Xueying, QIAN Cheng, et al. Feeding state classification of grass carp based on optical flow and improved RNN[J]. Acta Hydrobiologica Sinica, 2022, 46(6): 914 – 921. (in Chinese)

- [26] WEI D, BAO E, WEN Y, et al. Behavioral spatial-temporal characteristics-based appetite assessment for fish school in recirculating aquaculture systems[J]. *Aquaculture*, 2021, 545: 737215.
- [27] DONAHUE J, ANNE HENDRICKS L, GUADARRAMA S, et al. Long-term recurrent convolutional networks for visual recognition and description[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 2625 – 2634.
- [28] YUE-HEI N G J, HAUSKNECHT M, VIJAYANARASIMHAN S, et al. Beyond short snippets: deep networks for video classification[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 4694 – 4702.
- [29] SRIVASTAVA N, MANSIMOV E, SALAKHUDINOV R. Unsupervised learning of video representations using LSTMS[C] // *International Conference on Machine Learning*. PMLR, 2015: 843 – 852.
- [30] WU Z, WANG X, JIANG Y G, et al. Modeling spatial-temporal clues in a hybrid deep learning framework for video classification[C] // *Proceedings of the 23rd ACM International Conference on Multimedia*, 2015: 461 – 470.
- [31] 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)
- [32] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132 – 7141.
- [33] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting [C] // *Advances in Neural Information Processing Systems*, 2015: 802 – 810.
- [34] DU L, LU Z, LI D. Broodstock breeding behaviour recognition based on Resnet50 – LSTM with CBAM attention mechanism [J]. *Computers and Electronics in Agriculture*, 2022, 202: 107404.
- [35] 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.
- [36] SIMONYAN K, ZISSERMAN A. Two-stream convolutional networks for action recognition in videos[J]. *Advances in Neural Information Processing Systems*, 2014, 27: 568 – 576.
- [37] WANG L, XIONG Y, WANG Z, et al. Towards good practices for very deep two-stream convnets[J]. *arXiv preprint arXiv: 1507.02159*, 2015.
- [38] FEICHTENHOFER C, PINZ A, ZISSERMAN A. Convolutional two-stream network fusion for video action recognition[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1933 – 1941.
- [39] WANG L, XIONG Y, WANG Z, et al. Temporal segment networks: towards good practices for deep action recognition[C] // *European Conference on Computer Vision*. Springer, Cham, 2016: 20 – 36.
- [40] ZHOU B, ANDONIAN A, OLIVA A, et al. Temporal relational reasoning in videos [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 803 – 818.
- [41] FEICHTENHOFER C, FAN H, MALIK J, et al. SlowFast networks for video recognition[C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019: 6202 – 6211.
- [42] WANG G, MUHAMMAD A, LIU C, et al. Automatic recognition of fish behavior with a fusion of RGB and optical flow data based on deep learning[J]. *Animals*, 2021, 11(10): 2774.
- [43] DITRIA E M, JINKS E L, CONNOLLY R M. Automating the analysis of fish grazing behaviour from videos using image classification and optical flow[J]. *Animal Behaviour*, 2021, 177: 31 – 37.
- [44] CHÉRON G, LAPTEV I, SCHMID C. P-CNN: pose-based CNN features for action recognition[C] // *Proceedings of the IEEE International Conference on Computer Vision*, 2015: 3218 – 3226.
- [45] WU Z, FU Y, JIANG Y G, et al. Harnessing object and scene semantics for large-scale video understanding [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 3112 – 3121.
- [46] WANG Y, SONG J, WANG L, et al. Two-stream SR – CNNs for action recognition in videos[C] // *BMVC*, 2016.
- [47] WU Z, JIANG Y G, WANG X, et al. Multi-stream multi-class fusion of deep networks for video classification[C] // *Proceedings of the 24th ACM International Conference on Multimedia*, 2016: 791 – 800.
- [48] ZHU Y, NEWSAM S. Depth2Action: exploring embedded depth for large-scale action recognition[C] // *Computer Vision – ECCV 2016 Workshops*, 2016: 668 – 684.
- [49] WANG J H, LEE S K, LAI Y C, et al. Anomalous behaviors detection for underwater fish using AI techniques[J]. *IEEE Access*, 2020, 8: 224372 – 224382.
- [50] FOURNIER J C. *Théorie des graphes et applications: avec exercices et problèmes*[M]. Chippenham: Antony Rowe Ltd., Lavoisier, 2011.
- [51] BERNDT D J, CLIFFORD J. Using dynamic time warping to find patterns in time series[C] // *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 1994: 359 – 370.
- [52] CHENG S, ZHAO K, ZHANG D. Abnormal water quality monitoring based on visual sensing of three-dimensional motion behavior of fish[J]. *Symmetry*, 2019, 11(9): 1179.
- [53] 罗冰. 基于 PSO_BPNN 的鲫鱼浮头数量监测模型[D]. 北京: 中国农业大学, 2020.

LUO Bing. Monitoring model of crucian carp hypoxia based on PSO_BPNN[D]. Beijing: China Agricultural University, 2020. (in Chinese)

[54] 张笑华, 吴圣斌, 方圣恩, 等. 采用 Pareto 人工鱼群算法的结构健康监测传感器位置多目标优化[J]. 振动工程学报, 2022, 35(2): 351–358.

ZHANG Xiaohua, WU Shengbin, FANG Sheng'en, et al. Multi-objective sensor optimal placement for structural health monitoring based on Pareto artificial fish swarm algorithm[J]. Journal of Vibration Engineering, 2022, 35(2): 351–358. (in Chinese)

[55] COOK D, MIDDLEMISS K, JAKSONS P, et al. Validation of fish length estimations from a high frequency multi-beam sonar (ARIS) and its utilisation as a field-based measurement technique[J]. Fisheries Research, 2019, 218: 59–68.

[56] 罗毅, 王伟, 刘勇, 等. 基于 One-Class SVM 的青鲈鱼异常行为识别方法[J]. 河北工业科技, 2022, 39(3): 230–236.

LUO Yi, WANG Wei, LIU Yong, et al. Abnormal behavior recognition method of medaka based on One-Class SVM[J]. Hebei Journal of Industrial Science and Technology, 2022, 39(3): 230–236. (in Chinese)

[57] HSIAO Y H, CHEN C C, LIN S I, et al. Real-world underwater fish recognition and identification, using sparse representation [J]. Ecological Informatics, 2014, 23: 13–21.

[58] KALPATHY-CRAMER J, DE HERRERA A G S, DEMNER-FUSHMAN D, et al. Evaluating performance of biomedical image retrieval systems—an overview of the medical image retrieval task at ImageCLEF 2004—2013 [J]. Computerized Medical Imaging and Graphics, 2015, 39: 55–61.

[59] SALMAN A, JALAL A, SHAFAIT F, et al. Fish species classification in unconstrained underwater environments based on deep learning[J]. Limnology and Oceanography: Methods, 2016, 14(9): 570–585.

[60] JÄGER J, RODNER E, DENZLER J, et al. SeaCLEF 2016: object proposal classification for fish detection in underwater videos[C]//CLEF (Working Notes), 2016: 481–489.

[61] JÄGER J, WOLFF V, FRICKE-NEUDERTH K, et al. Visual fish tracking: combining a two-stage graph approach with CNN-features[C]//OCEANS 2017-Aberdeen. IEEE, 2017: 1–6.

[62] CUTTER G, STIERHOFF K, ZENG J. Automated detection of rockfish in unconstrained underwater videos using haar cascades and a new image dataset: labeled fishes in the wild[C]//2015 IEEE Winter Applications and Computer Vision Workshops. IEEE, 2015: 57–62.

[63] LI Z, LI W, LI F, et al. A review of computer vision technologies for fish tracking[J]. arXiv preprint arXiv:2110.02551, 2021.

[64] YANG L, LIU Y, YU H, et al. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: a review[J]. Archives of Computational Methods in Engineering, 2021, 28: 2785–2816.

[65] 董金耐, 杨森, 谢卓冉, 等. 水下图像目标检测数据集及检测算法综述[J]. 海洋技术学报, 2022, 41(5): 60–72.

DONG Jinnai, YANG Miao, XIE Zhuoran, et al. Overview of underwater image object detection data set and detection algorithms[J]. Journal of Ocean Technology, 2022, 41(5): 60–72. (in Chinese)

[66] LEVY D, BELFER Y, OSHEROV E, et al. Automated analysis of marine video with limited data[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018: 1385–1393.

[67] LIU L, LU H, CAO Z, et al. Counting fish in sonar images[C]//2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, 2018: 3189–3193.

[68] LI W, LI F, LI Z. CMFTNet: multiple fish tracking based on counterpoised JointNet[J]. Computers and Electronics in Agriculture, 2022, 198: 107018.