doi:10.6041/j.issn.1000-1298.2025.10.016

基于水质参数融合-多元 DeepAR 算法的蟹塘投饲时机判断方法

韩雨生[1]　王新猛[1]　王　浩[1]　胡佳宁[2]　任振辉[1]

(1.河北农业大学机电工程学院,保定 071001;2.石家庄鑫农机械有限公司,石家庄 052400)

摘要:我国蟹塘养殖行业大都采用固定时间投喂饲料,较少结合河蟹实际摄食强度,饲料利用率较低,蟹塘水质变化直接影响河蟹摄食强度,以此为基础,结合多元 DeepAR 算法,设计了一种蟹塘投饲时机的判断方法。选取溶解氧含量、水温、pH 值、氧化还原电位、电导率、溶解性总固体含量、氨氮含量和浊度 8 种水质参数为评价指标,选择三标度层次分析法为主观赋权法,基于熵权法的 CRITIC 法为客观赋权法,博弈论为组合赋权法,采用主客观组合赋权法构建了蟹塘水质评分方法;以时间序列、溶解氧含量和浊度为特征值,蟹塘水质分数为目标,利用多元 DeepAR 算法预测未来 12 h 水质分数曲线变化,曲线峰值是水质分数最大值,河蟹摄食强度等级最强,将曲线波峰阶段定义为最佳投饲时机。结果表明,蟹塘水质评分方法与实际河蟹摄食活动具有较好的关联性,多元 DeepAR 算法可以更好地预测未来 12 h 水质分数曲线变化,能够指导河蟹投饲环节。

关键词:水质参数融合;多元 DeepAR 算法;博弈论;河蟹摄食强度;蟹塘投饲时机

中图分类号:TP183;S966.16　　**文献标识码**:A　　**文章编号**:1000-1298(2025)10-0156-09　　**OSID**:

Method of Feeding Timing in Crab Ponds Based on Water Quality Parameter Fusion - Multivariate DeepAR Algorithm

HAN Yusheng[1]　WANG Xinmeng[1]　WANG Hao[1]　HU Jianing[2]　REN Zhenhui[1]

(1. *College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding* 071001, *China*
2. *Shijiazhuang Xinnong Machinery Co., Ltd., Shijiazhuang* 052400, *China*)

Abstract: In China's crab pond aquaculture industry, feed is mostly fed at fixed times, with little consideration of the actual feeding intensity of river crabs, resulting in low feed utilization efficiency. Changes in water quality in crab ponds directly affected the feeding intensity of river crabs. Based on this, combined with the multivariate DeepAR algorithm, a method for determining the optimal feeding time in crab ponds was designed. Eight water quality parameters, including dissolved oxygen content, water temperature, pH value, oxidation-reduction potential, electrical conductivity, total dissolved solids, ammonia nitrogen content, and turbidity, were selected as evaluation indicators. The three-scale analytic hierarchy process was chosen as the subjective weighting method, the CRITIC method based on the entropy weight method as the objective weighting method, and game theory as the combined weighting method. A crab pond water quality scoring method was constructed by using the subjective-objective combined weighting method. Taking time series, dissolved oxygen content, and turbidity as characteristic values, and the crab pond water quality score as the target, the multivariate DeepAR algorithm was used to predict the changes in the water quality score curve over the next 12 h. The peak of the curve represented the maximum water quality score, at which time the feeding intensity of river crabs was the highest. Therefore, the peak stage of the curve was defined as the optimal feeding time. The results showed that the crab pond water quality scoring method had a good correlation with the actual feeding activities of river crabs, and the multivariate DeepAR algorithm can better predict the changes in the water quality score curve over the next 12 h, which can guide the feeding process of river crabs.

Key words: fusion of water quality parameters; multivariate DeepAR algorithm; game theory; feeding intensity of river crabs; feeding timing in crab pond

收稿日期:2025-03-24　修回日期:2025-04-10

基金项目:河北省重点研发计划项目(20327217D)

作者简介:韩雨生(2000—),男,硕士生,主要从事河蟹养殖数智化研究,E-mail:hanyvsheng228@163.com

通信作者:任振辉(1968—),男,教授,博士,主要从事智能检测和自动控制技术研究,E-mail:renzh@hebau.edu.cn

0　引言

根据《2023 中国渔业年鉴》统计数据可知，我国河蟹养殖年产量约 8×10^5 t，总产值将近千亿元，已成为我国水产养殖中发展迅速、具有前景的支柱产业[1]。但我国蟹塘目前大都采用定时定量投饲，易造成投饲过量，水中残余饵料会加剧水质污染，影响蟹类生长速度和健康状况[2]。随着河蟹养殖行业的快速发展，研究精准投饲策略，提高饲料转化率是实现河蟹养殖行业绿色可持续发展的必经之路。

以水产动物摄食活动强度为依据投饲是水产养殖业的重要发展方向，近年来，采用机器视觉技术评估水产动物摄食强度的方法较多，ZHANG 等[3]为提高摄食强度识别精度，提出了一种基于双标签和 MobileViT－SENet 的池塘鱼群摄食强度识别方法；张铮等[4]针对主流的目标河蟹检测模型复杂度高，提出了基于 GhostNetV2 的 G3 模块以改进模型网络特征，并利用幻影卷积进行模型轻量化处理的方法；但目前河蟹主要是围网生态养殖、网箱养殖等开放式蟹塘养殖模式，机器视觉技术在此面临着水质混浊[5]、河蟹领地意识强、游泳能力差[6]等挑战，并不能很好地评估整个蟹群的摄食活动强度。WU 等[7]通过检测投喂时鱼群聚集和争斗对溶解氧含量的影响，开发了一种自适应的模糊投饲推理方法；BÓRQUEZ-LOPEZ 等[8]研究了溶解氧含量和水温对鱼类食欲的影响，结果表明，溶解氧含量对食欲影响较大（占 74%），水温影响较低（占 26%）；综上可知，溶解氧含量、水温等主要水质参数的变化会直接影响水产动物的食欲，水质传感器可监测开放式蟹塘整体水质随时间的变化情况[9]，可以根据水质变化评估食欲，预测并安排河蟹在最佳投饲时机摄食，增加河蟹日摄食量，提升河蟹年产量，提高饲料转化率，实现绿色养殖。

为此本文选取溶解氧含量、水温、pH 值、氧化还原电位（OPR）、电导率、溶解性总固体（TDS）含量、氨氮含量和浊度 8 种主要水质参数，采用主客观组合赋权的方法构建蟹塘水质评分方法，运用多元 DeepAR 时间序列算法预测未来蟹塘水质分数曲线变化，设定曲线波峰阶段为最佳投饲时机，蟹塘水质评分方法和多元 DeepAR 时间序列算法共同构成蟹塘投饲时机判断方法，以期为河蟹养殖行业的投饲环节提供信息技术支撑。

1　材料与方法

河蟹，学名中华绒螯蟹，属甲壳类的水生变温动物，据相关研究表明，河蟹摄食强度与多种水质参数相关，本文汇总分析了溶解氧含量、水温、pH 值、氧化还原电位（ORP）、电导率、溶解性总固体（TDS）含量、氨氮含量和浊度 8 种参数对河蟹摄食、生长具体影响情况，如表 1 所示[10-16]。

表 1　水质参数对河蟹生长具体影响

Tab.1　Influence of water parameters on growth of river crabs

参数	概念	作用
溶解氧含量	水中以气体形式溶解的氧气	溶解氧含量直接影响河蟹的消化吸收，水中溶解氧含量高，河蟹食量大、饵料效率高、生长快
水温	水体温度	水温直接影响河蟹生长、发育及新陈代谢；在一定范围内，随着水温增加，河蟹蜕壳速度加快，增重率增加，肥满度下降，摄食率增加
pH 值	水体酸碱度	在一定范围内，pH 值增加有利于幼蟹的蜕壳生长，超过一定范围时表现为抑制作用
ORP	衡量水体中氧化还原反应能力	判定水体污染水平，ORP 变化可能影响水中溶解物质的种类及浓度，从而影响河蟹的嗅觉和味觉感知，影响其对食物的发现和选择
电导率	电流在水体中传导能力	电导率过高可能意味着盐分或污染物浓度过高，而过低则可能表示水质较为稀薄，河蟹的产量和质量与其存在显著负相关
TDS	水体溶解的所有固体物质总量	影响水体渗透压，进而影响河蟹体内水分平衡，过高的 TDS 可能导致河蟹脱水，过低的 TDS 可能影响其营养物质的吸收
氨氮含量	水体中以游离氨（NH_3）和离子氨（NH_4^+）形式存在的氮	氨氮含量可以直接或间接影响养殖河蟹的生长和繁殖
浊度	水体中有机物、无机物、微生物等对光线的散射和吸收能力	浊度高，水质不良，会导致河蟹白天上岸活动频繁，口吐泡沫，生长减慢，成活率低

图 1 为蟹塘投饲时机方法框架，在蟹塘水质评分方法中，利用蟹塘水质变化映射河蟹摄食强度，首先利用三标度层次分析法和基于熵权法的 CRITIC 法分别计算 8 个指标权重，层次分析法为主观赋权法，将专家组的养殖经验融入蟹塘水质评分方法中，避免评价结果受到数据的随机误差影响；CRITIC 法是客观赋权法，反映水质参数的变异特征，有效地预防了加权人为的主观性；其次利用博弈论计算主客观的组合权重，设为最优权重，以弥补单一赋权带来的不足，消除误差，提高评估方法合理性；最后对 8

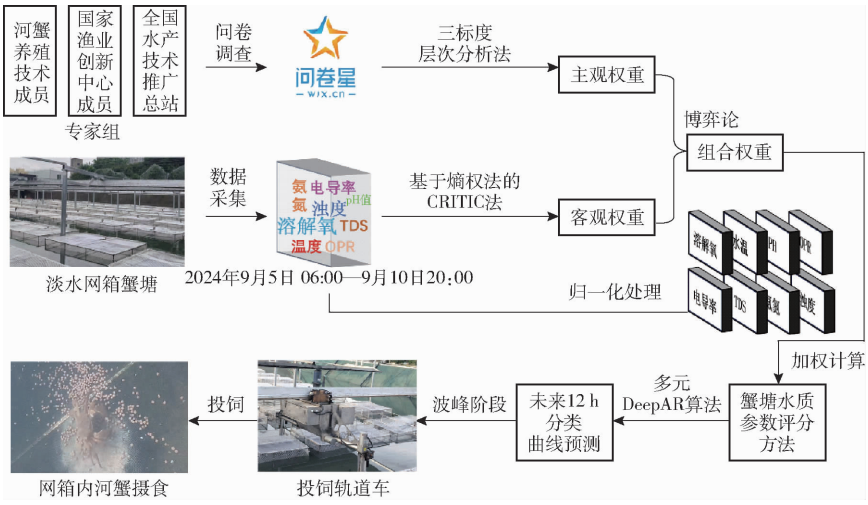

图 1 蟹塘投饲时机方法框架图

Fig. 1 Framework diagram of crab pond feeding timing method

个指标数据进行归一化处理,将其值乘以其对应的权重,并将所有加权得分相加,得到蟹塘水质评估分数。采用多元 DeepAR 算法预测未来 12 h 蟹塘水质分数的变化,寻找未来 12 h 蟹塘水质分数的波峰阶段,规定此阶段为最佳投饲时机,指导投饲作业。

2 蟹塘水质评分方法

2.1 主观权重计算

层次分析法是基于人类实践理论和丰富经验判断出参数重要性的主观赋权法。本文采用的层次分析法,将传统的“九标度法”改进为“三标度法”,降低了判断不同参数之间重要性的难度;将原先计算中的一致性检验步骤改进为构造一致化矩阵,省略了一致性检验可能造成的繁琐计算。具体步骤如下[17]:

(1)邀请从事河蟹养殖技术成员 2 名、国家渔业创新中心成员 2 名、全国水产技术推广总站成员 1 名,共 5 位组成专家组,采用三标度法对 8 种评价指标两两进行比较,得出比较矩阵(A1 ~ A5)。

(2)一致化矩阵元素 D_{ij} 计算式为

$$D_{ij} = \mathrm{e}^{\frac{A1_{ij} + A2_{ij} + A3_{ij} + A4_{ij} + A5_{ij}}{5}} \tag{1}$$

(3)计算一致化矩阵 $\boldsymbol{D}_{ij}$ 的特征向量,通过计算得出的特征向量集,即为主观赋权法所得参数权重 W_j'

$$W_j' = \left(\prod_{i=1}^{n} d_{ki} \right)^{\frac{1}{n}} \tag{2}$$

式中 d_{ki}——矩阵 $\boldsymbol{D}_{ij}$ 的列变量 n——行数

2.2 客观权重计算

客观权重由水质数据集的熵值、波动性和冲突性构成。熵值可以量化水质参数的无序性和变化幅度,若熵值较高,则表明水质参数随着时间序列变化剧烈;熵值较低,则表明水质参数随着时间序列变化平缓。不同的水质参数会表现出不同的波动性,水质参数会随着时间变化表现出一定的波动,是水质参数数据集的客观固有性质,水质参数的波动会直接影响河蟹摄食强度,例如:白天在阳光照射下蟹塘内、周边植物会经过光合作用增加蟹塘水体氧气含量,夜晚植物、动物和微生物呼吸会大量消耗蟹塘水体氧气,造成溶解氧含量在昼夜之间显著波动,为此,在清晨河蟹会因溶解氧含量不足而浮出水面,摄食强度低,在傍晚因溶解氧含量充足而食性大,摄食强度高;水温也会随着昼夜、季节等时间量纲变化而变化,水温影响河蟹脱壳,脱壳前能量集中于脱壳行为,消化系统能量供给较少,食欲较差,脱壳后,身体需要快速补充能量,食欲强,河蟹快速生长。冲突性指水质参数间相互作用,例如:水温与溶解氧含量呈负相关,温度越高,溶解氧含量越低;pH 值越低,溶解氧含量越低。

在江苏省宜兴市高塍镇物联网水产养殖基地的淡水网箱的蟹塘中部署了溶解氧含量、水温、pH 值、OPR、电导率、TDS、氨氮含量和浊度 8 种水质传感器,采集 2024 年 9 月 5 日 18:00—9 月 6 日 06:00 的数据,每隔 1 min 采集一次,每种水质参数均采集 720 条数据,组成 720 行 8 列数据矩阵,计算数据矩阵的熵、波动性和冲突性。因篇幅有限,仅展示每小时采集到的原始数据,如表 2 所示。

熵权法计算熵值[18],CRITIC 法计算波动性和冲突性[19],步骤如下所示:

(1)根据评价指标建立评价矩阵,设评价矩阵 $\boldsymbol{X}_{ij}$,i 组评价对象,j 种评价指标;对评价矩阵用最大最小值归一化公式进行无量纲化处理,得到矩阵 $\boldsymbol{X}_{ij}^{*}$;在本文数据集中,i 为 720,j 为 8。

(2)计算熵值 ewm_j 可以体现出参数间离散程度,计算式为

表 2　采集的部分原始数据

Tab. 2　Part of raw data collected

时刻	溶解氧含量/(mg·L^{-1})	水温/℃	pH 值	ORP/mV	电导率/(μS·cm^{-1})	TDS/(mg·L^{-1})	氨氮含量/(mg·L^{-1})	浊度/NTU
18:00	5.02	25.0	7.87	56	22.9	14.6	0.24	2.98
19:00	5.14	24.8	7.85	46	22.7	14.5	0.22	4.10
20:00	5.35	24.6	7.87	47	24.0	15.3	0.23	4.98
21:00	5.22	24.3	7.90	59	24.2	15.4	0.24	5.50
22:00	5.25	23.8	7.86	30	24.2	15.4	0.23	6.17
23:00	5.14	23.9	7.84	55	24.5	15.6	0.24	6.41
24:00	5.07	24.4	7.84	61	24.5	15.6	0.25	6.49
01:00	4.98	24.6	7.79	78	24.5	15.6	0.25	6.41
02:00	4.86	24.9	7.78	59	24.5	15.6	0.25	9.20
03:00	4.49	25.0	7.78	79	24.5	15.6	0.25	9.03
04:00	4.36	25.2	7.78	83	24.3	15.5	0.26	8.83
05:00	4.43	25.4	7.79	75	24.3	15.5	0.27	8.53
06:00	4.31	25.6	7.78	80	24.9	15.9	0.27	8.96

$$\mathrm{ewm}_j = \frac{1 - e_j}{\sum_{j=1}^{i}(1 - e_j)} \tag{3}$$

其中
$$\mathrm{e}_j = -\frac{1}{ln n}\sum_{i=1}^{n}\frac{\boldsymbol{X}_{ij}^*}{\sum_{i=1}^{n}\boldsymbol{X}_{ij}^*}\,ln\,\frac{\boldsymbol{X}_{ij}^*}{\sum_{i=1}^{n}\boldsymbol{X}_{ij}^*} \tag{4}$$

式中　e_j——第 j 项参数的信息熵值

(3)参数波动性 S_j 计算式为

$$\mathrm{S}_j = \sqrt{\frac{\sum_{i=1}^{n}(\boldsymbol{X}_{ij}^* - \overline{\boldsymbol{X}}_j^*)^2}{n-1}} \tag{5}$$

式中　$\overline{\boldsymbol{X}}_j^*$——在矩阵 $\boldsymbol{X}^*$ 中 8 个参数的列均值

(4)参数间冲突性 f_j 计算式为

$$f_j = \sum_{i=1}^{m}(1 - |r_{ij}|) \tag{6}$$

式中　r_{ij}——皮尔逊相关系数[20]

(5)将熵值、波动性和冲突性代入到客观权重表达式，W_j'' 为客观权重

$$W_j'' = \frac{(ewm_j + S_j)f_j}{\sum_{j=1}^{n}(ewm_j + S_j)\sum_{j=1}^{n}f_j} \tag{7}$$

2.3　组合权重计算

在河蟹水质评分方法中，仅使用主观权重加权计算水质分数存在缺乏科学性与客观性依据、易受人为偏见和主观判断影响等问题；仅使用客观权重存在受数据质量限制和影响、忽视因子间相互作用和专家经验等问题。组合权重是将主客观权重结合，充分发挥主观和客观权重的优势，弥补单一方法的不足之处，提升评估方法的灵活性、提高模型的可解释性。

采用博弈论组合赋权法进行主客观权重融合，融合思想是寻求主观权重和客观权重集合的最小化离差，使水质分数加权计算结果更具有合理性，根据博弈论模型基本权重之间各自偏差的原理，寻找最满意的权向量，优化线性组合系数 α_k[21]，为

$$\begin{bmatrix}\sum_{j=1}^{8}(W_j')^2 & \sum_{j=1}^{8}W_j'W_j'' \\ \sum_{j=1}^{8}W_j'W_j'' & \sum_{j=1}^{8}(W_j'')^2\end{bmatrix}\begin{bmatrix}\alpha_1 \\ \alpha_2\end{bmatrix} = \begin{bmatrix}\sum_{j=1}^{8}(W_j')^2 \\ \sum_{j=1}^{8}(W_j'')^2\end{bmatrix} \tag{8}$$

计算得出 α_1、α_2 后，使用最大最小值归一化公式后得到 α_1^*、α_2^*，即

$$W_j''' = \alpha_1^* W_j' + \alpha_2^* W_j'' \tag{9}$$

式中　W_j'''——基于博弈论的组合权重

表 3 为 3 种权重计算结果汇总。

表 3　3 种权重计算结果汇总

Tab. 3　Results of summarized three weight calculations

指标	主观权重	客观权重	组合权重
溶解氧含量	0.222	0.091	0.170
水温	0.169	0.148	0.160
pH 值	0.165	0.110	0.142
ORP	0.064	0.123	0.088
电导率	0.064	0.133	0.092
TDS	0.063	0.134	0.092
氨氮含量	0.159	0.192	0.172
浊度	0.094	0.070	0.084

2.4　水质分数加权计算

评价指标之间量纲差异，最大最小值不相同，无

法进行加权计算，为此，设定水质评分标准为百分制，水质评分最大值为 100 分；8 项评价指标，水质评分最大值除以 8；可得每种评价指标最大值是 12.5 分。

故设每种评价指标最大值是 12.5 分，最小值是 0 分，采用归一化公式进行无量纲处理，使评价指标数据集合位于区间(0，12.5)，即

$$x' = \frac{12.5(x - \min(x))}{\max(x) - \min(x)} \tag{10}$$

式中 x'——归一化后数据

x——原始数据

$\min(x)$——原数据最小值

$\max(x)$——原数据最大值

归一化后各水质参数分别乘以对应权重，将所得值累加，构成蟹塘水质分数。

2.5 蟹塘水质参数权重验证

为验证 3 种权重的合理性，从试验基地内选取 5 个网箱(A～E)，网箱内设置摄食区，采用相机记录摄食区河蟹活动，记录时间为 2024 年 9 月 5 日 06:00—9 月 6 日 06:00，共 24 h，5 min 记录一次，5 个网箱摄食区如图 2 所示。

通过人为记录 5 个网箱摄食区内河蟹总数，由于河蟹具有领地意识，所以摄食区内不会同时存在较多河蟹，将蟹群摄食强度分为“微”、“弱”、“中”、“强”，分别对应摄食区河蟹总数为 0～3、4～7、8～11、12～20[22]。

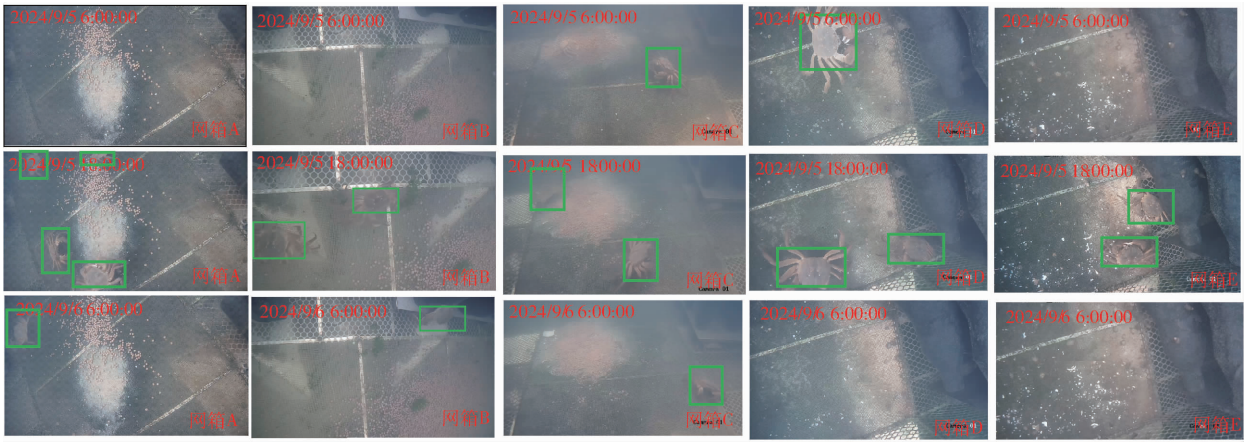

图 2 网箱摄食区

Fig. 2 Cage feeding area

2024 年 9 月 5 日 06:00—9 月 6 日 06:00 蟹塘水质分数和蟹群摄食强度等级如图 3 所示，可以看出，在 9 月 5 日 15:00—22:00 区域，蟹塘水质分数较高，蟹群摄食强度等级保持在中和强之间，与蟹塘水质分数相关，与河蟹昼伏夜出生活习性相符；在 9 月 5 日 06:00—8:00 和 9 月 6 日 04:00—06:00 区域，蟹塘水质分数较低，蟹群摄食强度等级保持在微和弱之间，在试验现场观察到蟹群分布在网箱上方，在底部摄食区的河蟹只数较少，水质分数与实际情况相符。

主观权重受人为偏见和主观判断影响造成水质分数偏高，客观权重受数据质量限制和影响造成水质分数偏低；组合权重充分融合了主观和客观权重优势，弥补单一方法的不足，使水质分数变化更柔和，位于两者中间，故选用组合权重构建蟹塘水质评分方法。表明本文构建的蟹塘水质评分方法与实际河蟹摄食活动具有较好的关联性，能够指导河蟹投饲环节。

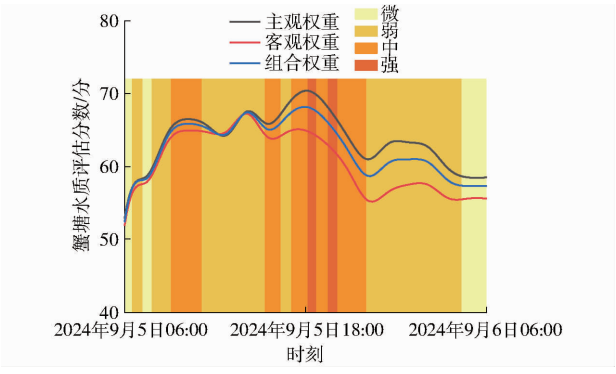

图 3 水质评估分数和蟹群摄食强度等级对比

Fig. 3 Comparison of water quality fraction and feeding intensity level of crab population

3 蟹塘水质分数时间序列预测算法

3.1 多元 DeepAR 算法模型

将未来 12 h 内水质分数曲线的峰值定义为最佳投饲时机，因此需要预测未来 12 h 水质分数随时间序列的变化。DeepAR 深度自回归循环神经网络(Deep learning probabilistic forecasting with autoregressive recurrent networks)算法是将深度学习与概率预测结合，在递归神经网络(Recurrent neural

network，RNN)基础上根据时间序列的过去生成时间序列未来的概率预测值监督学习算法。传统的 DeepAR 算法仅选用时间序列为特征值，缺少对有效特征的挖掘，本文在此基础上提出了多元 DeepAR 算法，因蟹塘水质分数由水质参数构成，可以在水质参数中选择与算法预测精度正相关的有效特征，参与算法预测，提高模型精度。本文构建的多元 DeepAR 模型如图 4 所示[23]。

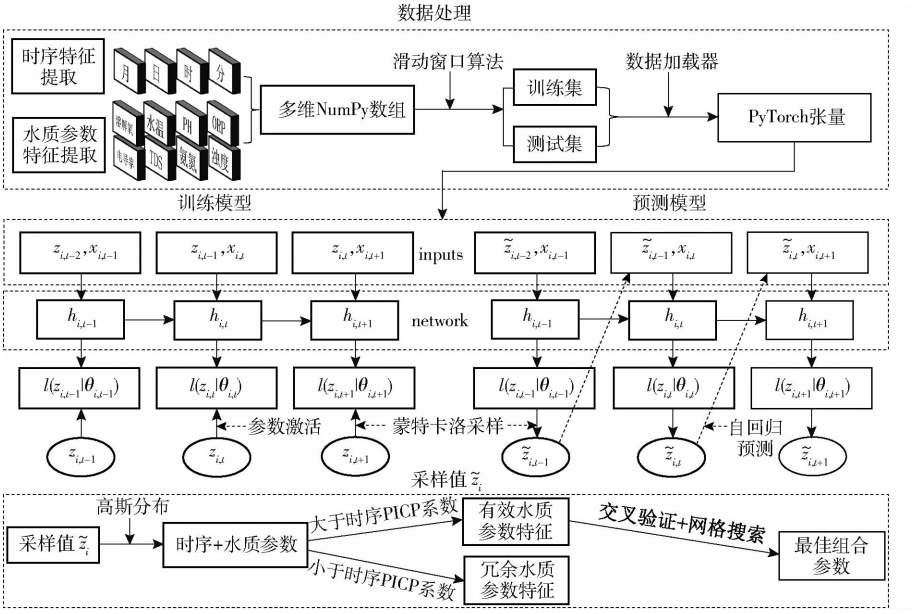

图 4　DeepAR 模型构建

Fig. 4　DeepAR model construction

(1)数据处理

从时间序列中提取月、日、时和分 4 种时序特征，在数据集中按照列名提取水质参数特征，以蟹塘水质分数为目标，构建多维 NumPy 数组；采用滑动窗口算法从输入的多维数组中提取数据，划分为训练集和测试集；构造数据加载器，将 NumPy 数组转换为 PyTorch 张量，以便于训练深度学习模型。

(2)训练模型

$z_{i,t}$为第 i 个序列在 t 时刻的观测值，$x_{i,t}$为第 i 个序列在 t 时刻的协变量。训练时在时间序列 t 上输入协变量、上一时刻观测值 $z_{i,t-1}$和上一时刻的输出 $h_{i,t-1}$，经过 RNN 循环神经网络，得到 $h_{i,t}=H(h_{t-1}, z_{t-1}, \boldsymbol{\Theta})$，此时通过 $z_{i,t}$与 $h_{i,t}$最大化对数似然函数，寻找最佳参数权重，对数似然函数的计算公式为

$$L=\sum_i \sum_t \lg l\{z_{i,t}[\theta(H(h_{t-1}, z_{t-1}, \boldsymbol{\Theta}))]\} \tag{11}$$

式中　$l\{z_{i,t}[\theta(h_{i,t})]\}$——似然函数

θ——似然参数

H——RNN 神经元内部函数

$\boldsymbol{\Theta}$——模型参数　　h_t——隐藏层输出

(3)预测模型

预测模型与训练模型结构相同，共享权重，在每个时刻 t 使用蒙特卡洛采样得到符合 $l(\cdot|\theta_t)$的采样值 $\tilde{z}_t$，将采样值 $\tilde{z}_t$ 输入下一个时间步长，进行自回归预测得到一系列采样值，利用采样值可以计算期望；本文采用高斯分布作为似然模型进行蟹塘水质分数区间预测，激活函数选取 softplus；选择置信度为 10% ~90% 的区间为预测区间、置信度为 50% 的区间预测作为水质环境分数的预测结果。

(4)时序 + 水质参数区间预测

近些年在 DeepAR 算法应用中均仅以时序为特征[24-25]，但神经网络模型的性能在很大程度上取决于输入特征的质量和数量，添加有效特征可以帮助模型更好地捕捉数据中的潜在模式和结构，从而提升预测的准确性。蟹塘水质评估方法由水质参数构建而成，为此添加水质参数为特征，提高模型精度。为筛选有效特征，进行时序 + 水质参数区间预测，为评价预测效果，引入预测区间覆盖率 PICP 系数，用于衡量预测区间包含真实值的概率，其值越大，意味着落在预测区间的真实值越多，计算公式为

$$PICP=\frac{1}{N}\sum_{i=1}^{N} I(y_i \in [L_i, U_i]) \tag{12}$$

式中 N 为样本总数，y_i 为第 i 个观测值，L_i、U_i 是分别为针对第 i 个观测值的预测区间的下界和上界，$I(y_i \in [L_i, U_i])$为指示函数，当 y_i 落在预测区间 $[L_i, U_i]$内时取值为 1，否则为 0。因此时序 + 水质参数区间预测模型大于时序区间预测模型 PICP 系数的水质参数为有效特征，小于为冗余特征。

(5)最佳参数组合

搭建含有两层 RNN 神经元隐藏层的 DeepAR 模型,优化器选择 Adam;最佳参数组合:学习率为 0.01,批大小为 32,隐藏层单元数为 40,时间步长为 5,Dropout 参数为 0.1,模型迭代次数为 100。

(6)时序 + 水质参数区间预测结果

如图 5 所示,图 5a 为特征变量仅为时序区间预测结果,预测区间覆盖率 PICP 系数 0.52;图 5b 为特征变量为时序 + 溶解氧含量区间预测结果,预测区间覆盖率 PICP 系数 0.72,置信度 50% 曲线与实际值曲线重合度较高,判定溶解氧含量为有效特征;图 5c 为特征变量为时序 + 水温区间预测结果,预测区间覆盖率 PICP 系数 0.08,判定水温为冗余特征;图 5d 为特征变量为时序 + pH 值区间预测结果,预测区间覆盖率 PICP 系数 0.30,判定 pH 值为冗余特征;图 5e 为特征变量为时序 + 电导率区间预测结果,预测区间覆盖率 PICP 系数 0.35,判定电导率为冗余特征;图 5f 为特征变量为时序 + 氨氮含量区间预测结果,预测区间覆盖率 PICP 系数 0.35,判定氨氮含量为冗余特征;图 5g 为特征变量为时序 + 浊度区间预测结果,预测区间覆盖率 PICP 系数 0.57,判定浊度为有效特征;图 5h 为特征变量为时序 + ORP 区间预测结果,预测区间覆盖率 PICP 系数 0.43,判定 ORP 为冗余特征;图 5i 为特征变量为时序 + TDS 区间预测结果,预测区间覆盖率 PICP 系数 0.43,判定 TDS 为冗余特征;根据上述分析开展多元 DeepAR 算法模型构建。

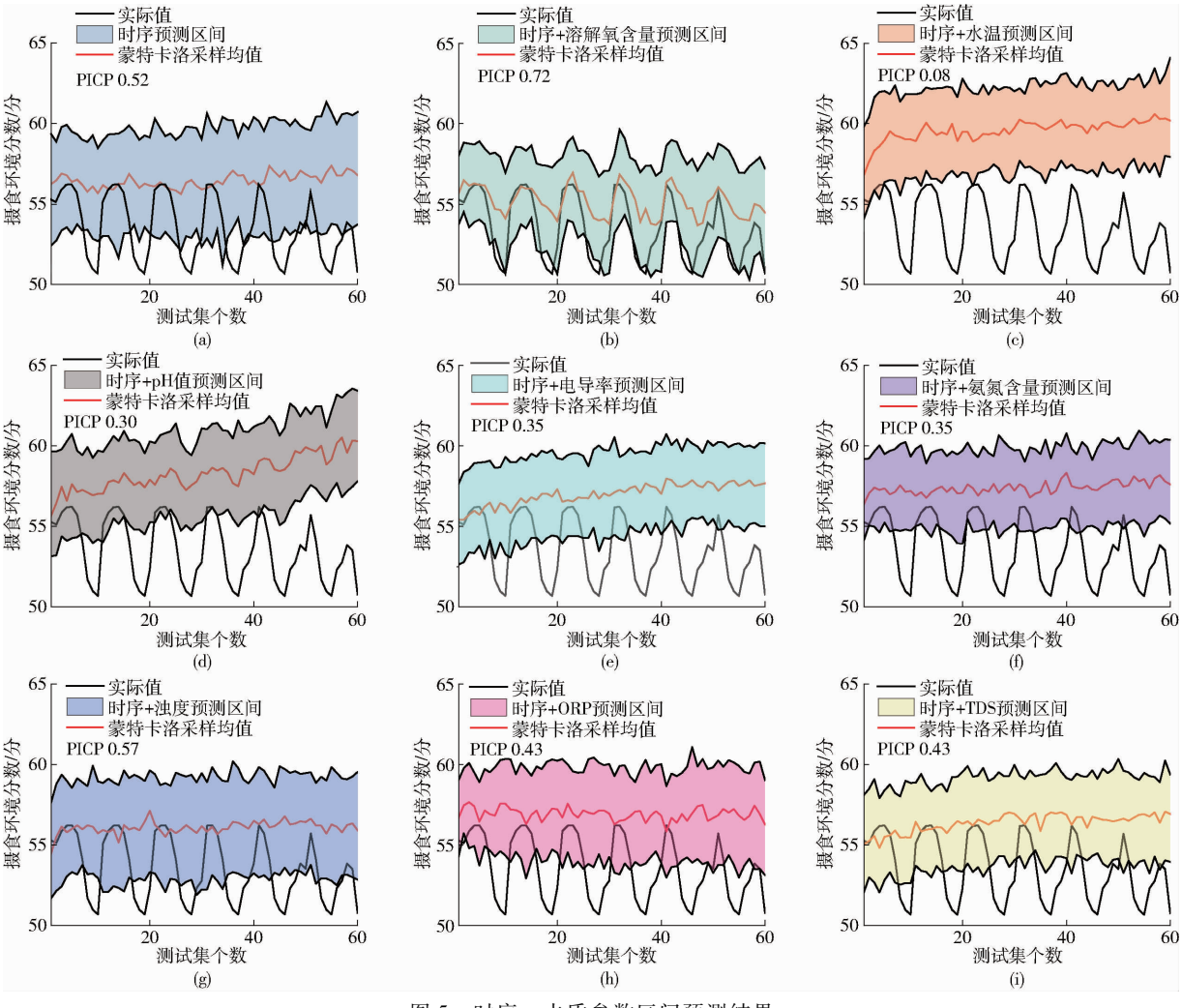

图 5 时序 + 水质参数区间预测结果

Fig. 5 Time series + water quality parameter interval prediction results

3.2 不同模型对比试验

选用预测模型性能评价指标有:均方误差(MSE)、平均绝对误差(MAE)和决定系数(R^2)。MSE 为预测值与实际值之差平方的期望值;MAE 为绝对误差的平均值,能反映预测值误差的实际情况;R^2 为将预测值与仅使用均值的情况相比,结果越接近 1,模型准确度越高。

为验证多元 DeepAR 算法预测性能,选取不同的特征值和不同的预测算法进行比较,特征值和预测算法共组合成 6 种不同模型,如表 4 所示。在特征值中分成 3 种组合,分别是输入特征值时序;输入特征值时序和有效特征(溶解氧含量、浊度);输入

表 4　特征值和算法组合模型

Tab. 4　Eigenvalues and algorithms combine models

模型序号	算法	输入该特征值								
		时序	溶解氧含量	水温	pH 值	电导率	氨氮含量	浊度	ORP	TDS
1	XGBoost	是	是	否	否	否	否	是	否	否
2	GRU	是	是	否	否	否	否	是	否	否
3	LSTM	是	是	否	否	否	否	是	否	否
4	DeepAR	是	是	否	否	否	否	是	否	否
5	DeepAR	是	否	否	否	否	否	否	否	否
6	DeepAR	是	是	是	是	是	是	是	是	是

特征值时序和全部水质参数(溶解氧含量、水温、pH 值、电导率、氨氮含量、浊度、ORP 和 TDS)。在预测算法方面,选取极端梯度提升(XGBoost)算法、门控循环单元(GRU)算法、长短期记忆网络(LSTM)算法和 DeepAR 算法,其中 GRU 算法和 LSTM 算法是常用的时间序列预测算法,已被广泛用于水质参数预测。

图 6 为 6 种模型在 2024 年 9 月 10 日 08:00—20:00 蟹塘水质分数预测曲线,可以看出 DeepAR 算法,特征值是时序、溶解氧含量和浊度模型整体预测曲线与真实值曲线最为接近,尤其是在波峰波谷处预测数值接近效果较为明显,拟合效果优于其他 5 种模型。

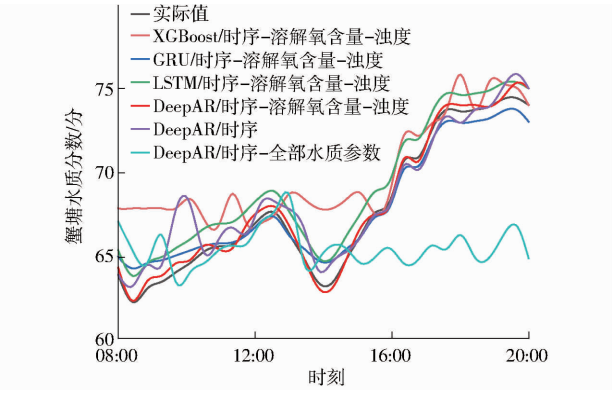

图 6　不同模型预测结果

Fig. 6　Different models predicting outcomes

表 5 为各模型在 2024 年 9 月 10 日 08:00—20:00 模型预测结果评价指标。从表 5 可以看出,使用多元 DeepAR 模型进行预测的 MSE、MAE 最小,R^2 最接近 1,表明多元 DeepAR 模型在预测未来 12 h 蟹塘水质分数时有更高的预测精度。

表 5　各模型预测结果误差分析

Tab. 5　Error analysis of prediction results of each model

模型序号	MSE/分^2	MAE/分	R^2
1	4. 591	1. 221	0. 890
2	2. 254	0. 884	0. 946
3	0. 194	0. 366	0. 971
4	0. 125	0. 295	0. 987
5	1. 543	0. 865	0. 907
6	22. 708	3. 525	0. 335

通过图 6 和表 5 可得:由模型 1、2、3 和 4 对比可知,算法预测效果从大到小依次为多元 DeepAR、LSTM、GRU、XGBoost,表明 DeepAR 算法优于其他传统时间序列预测算法,其中多元 DeepAR、LSTM 和 GRU 均是以神经网络算法为基础构建的,说明神经网络算法在水质数据预测方面表现较为优越,验证了通过选取有效特征提高 DeepAR 算法的预测精确方法的有效性;由模型 4、5、6 对比可知,模型 4 的 MSE、MAE 最小,R^2 最接近 1,预测数值最接近实际值,模型 5 预测曲线效果与实际值也表现出一定的重合度,说明 DeepAR 算法本身就具有较强的预测能力,模型 6 预测曲线预测效果最差,说明冗余特征值对预测效果影响较大,在 DeepAR 算法中添加有效特征会增强算法的预测结果,添加冗余特征会极大程度干扰算法的预测效果。

4　结论

(1)为寻找开放式淡水蟹塘河蟹养殖最佳投饲时机,提出了一种基于水质参数融合-多元 DeepAR 算法的蟹塘投饲时机判断方法。

(2)在蟹塘水质评估方法中,选择溶解氧含量、水温、pH 值、氧化还原电位(OPR)、电导率、溶解性总固体(TDS)含量、氨氮含量和浊度 8 种水质参数为评价指标,运用三标度层次分析法计算主观权重,基于熵权法的 CRITIC 法计算客观权重,为兼顾主客观权重优点,利用博弈论法计算了组合权重,得到溶解氧含量权重为 0. 170、水温权重为 0. 160、pH 值权重为 0. 142、氧化还原电位权重为 0. 088、电导率权重为 0. 092、溶解性总固体含量权重为 0. 092、氨氮含量权重为 0. 172、浊度权重为 0. 084;通过蟹塘水质分数变化与摄食区河蟹只数关联性试验分析验证了水质参数权重的合理性。

(3)在蟹塘水质分数时间序列预测模型中,以月、日、时、分、溶解氧含量和浊度为特征,蟹塘水质

评估分数为目标,训练好的多元 DeepAR 模型均方误差(MSE)为 0.125 分^2、平均绝对误差(MAE)为 0.295 分、决定系数(R^2)为 0.987,优于其他常规时间序列预测算法,可以较好预测蟹塘水质分数未来 12 h 的变化曲线,能够指导河蟹投饲环节,促进河蟹健康生长。

参 考 文 献

[1] 农业农村部渔业渔政管理局, 全国水产技术推广总站, 中国水产学会. 2023 中国渔业统计年鉴[M]. 北京:中国农业出版社,2023.

[2] 雷高辉,刘峰,董小宁,等. 水产养殖智能投饵技术研究进展[J]. 饲料工业,2024,45(14):132 – 144.
LEI Gaohui, LIU Feng, DONG Xiaoning, et al. Research progress of intelligent feeding technology in aquaculture[J]. Feed Industry,2024,45(14):132 – 144. (in Chinese)

[3] ZHANG Lu, LIU Zunxu, ZHANG Yapeng, et al. Feeding intensity identification method for pond fish school using dual-label and MobileViT-SENet[J]. Biosystems Engineering,2024,241:113 – 128.

[4] 张铮,鲁祥,胡庆松. 基于图像增强与 GC – YOLO v5s 的水下环境河蟹识别轻量化模型研究[J]. 农业机械学报,2024,55(11):124 – 131,374.
ZHANG Zheng, LU Xiang, HU Qingsong. Lightweight model for river crab detection based on image enhancement and improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery ,2024,55(11):124 – 131,374. (in Chinese)

[5] 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)

[6] 胡庆松,朱皓,李俊. 虾蟹养殖池塘机械化研究现状及发展趋势[J]. 上海海洋大学学报,2022,31(5):1216 – 1223.
HU Qingsong, ZHU Hao, LI Jun. Research progress and development trend of mechanization of shrimp and crab breeding pond [J]. Journal of Shanghai Ocean University, 2022,31(5):1216 – 1223. (in Chinese)

[7] WU T H, HUANG Y I, CHEN J M. Development of an adaptive neural-based fuzzy inference system for feeding decision-making assessment in silver perch (*Bidyanus bidyanus*) culture[J]. Aquacultural Engineering, 2015,66: 41 – 51.

[8] BÓRQUEZ-LOPEZ R A, CASILLAS-HERNANDEZ R, LOPEZ ELIAS J A, et al. Improving feeding strategies for shrimp farming using fuzzy logic , based on water quality parameters[J]. Aquacultural Engineering, 2018,81: 38 – 45.

[9] 樊宇星,任妮,田港陆,等. 基于 DeepAR-RELM 的池塘溶解氧含量时空预测方法研究[J]. 农业机械学报,2020,51(增刊 1):405 – 412.
FAN Yuxing, REN Ni, TIAN Ganglu, et al. Spatio-temporal prediction method of dissolved oxygen in ponds based on DeepAR-RELM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(Supp. 1): 405 – 412. (in Chinese)

[10] 李多喜,丁传珍,王黎明. 浅谈水环境与河蟹摄食[J]. 水产科技情报,1988(3):12 – 13.

[11] 张凯军,王静安,倪康达,等. 不同水温对个体养殖方法中华绒螯蟹幼蟹蜕壳生长和摄食率的影响[J]. 淡水渔业,2021,51(5):100 – 105.

[12] 刘金生,王军,岳武成,等. 水体 pH 值对中华绒螯蟹幼蟹蜕壳生长及其相关基因表达的影响[J]. 淡水渔业,2016,46(4):96 – 100.

[13] 李春波,沈晨晨,冯广朋,等. 中华绒螯蟹栖息地因子适宜范围的研究进展[J]. 海洋渔业,2023,45(6):766 – 774.
LI Chunbo, SHEN Chenchen, FENG Guangpeng, et al. Research progress on habitat suitability of *Eriocheir sinensis*[J]. Marine Fisheries, 2023,45(6): 766 – 774. (in Chinese)

[14] 彭智文,黄姝,岳武成,等. 不同盐度下中华绒螯蟹胁迫相关基因的表达差异[J]. 福建农林大学学报(自然科学版),2017,46(5):552 – 556.
PENG Zhiwen, HUANG Shu, YUE Wucheng, et al. Effect of different salinity on the expression profiles of stress-relative genes in Chinese mitten crab, *Eriocheir sinensis*[J]. Journal of Fujian Agriculture and Forestry University (Natural Science Edition), 2017,46(5): 552 – 556. (in Chinese)

[15] YANG Hu, LIU Jiang, WU Ming. Influence of water quality on the growth and feeding of *Eriocheir sinensis* in aquaculture [J]. Journal of Applied Aquaculture, 2018,30(3): 261 – 276.

[16] LI Zhi, WANG Xiao, WANG Jun. Impact of total dissolved solids on the growth and survival of the Chinese mitten crab, *Eriocheir sinensis*[J]. Journal of Shellfish Research, 2017,36(1): 87 – 93.

[17] 林文捷,赵明岩,陈浩克,等. 基于冷链物流的鲜鸡蛋抗压特性试验及参数优化[J]. 农业工程学报,2024,40(17):305 – 313.
LIN Wenjie, ZHAO Mingyan, CHEN Haoke, et al. Optimizing the parameters for compressive resistances of fresh eggs using cold chain logistics[J]. Transactions of the CSAE, 2024,40(17): 305 – 313. (in Chinese)

[18] 赵经华,马世骄,房城泰. 基于熵权-模糊综合评价法的无人机多光谱春玉米长势监测模型研究[J]. 农业机械学报,2024,55(8):214 – 224.
ZHAO Jinghua, MA Shijiao, FANG Chengtai. Growth monitoring of spring maize using UAV multispectral imaging based on entropy weight – fuzzy comprehensive evaluation method [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(8): 214 – 224. (in Chinese)

(下转第 183 页)

[21] VERMA H, KUMAR S. An accurate missing data prediction method using LSTM based deep learning for health care[C]// Proceedings of the 20th International Conference on Distributed Computing and Networking, 2019: 371 – 376.
[22] ZHANG Y, THORBURN P, XIANG W, et al. SSIM—a deep learning approach for recovering missing time series sensor data [J]. IEEE Internet of Things Journal, 2019, 6(4): 6618 – 6628.
[23] AHMED S, NIELSEN I E, TRIPATHI A, et al. Transformers in time-series analysis: a tutorial[J]. Circuits, Systems, and Signal Processing, 2023, 42(12): 7433 – 7466.
[24] 陈红, 夏青, 左婷, 等. 基于机器视觉的花菇分选技术[J]. 农业机械学报, 2014, 45(1): 281 – 287.
CHEN Hong, XIA Qing, ZUO Ting, et al. Determination of shiitake mushroom grading based machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(1): 281 – 287. (in Chinese)
[25] ARCHITH S, YUKTA C, ARCHANA H, et al. Analysis of M – SEIR and LSTM models for the prediction of COVID – 19 using RMSLE[M]// SINGH P. Fundamentals and methods of machine and deep learning: algorithms, tools and applications. Hoboken: John Wiley & Sons, 2022: 101 – 119.
[26] WANNINKHOF R, LEDWELL J R, BROECKER W S. GAS exchange-wind speed relation measured with sulfur hexafluoride on a lake[J]. Science, 1985, 227(4691): 1224 – 1226.
[27] GALLAGHER B K, GEARGEOURA S, FRASER D J. Effects of climate on salmonid productivity: a global meta-analysis across freshwater ecosystems[J]. Global Change Biology, 2022, 28(24): 7250 – 7269.
[28] 鲁植雄, 王雨彤, 王琳,等. 基于 IGWPSO – SVM 的 HMCVT 湿式离合器摩擦副温度预测[J]. 农业机械学报, 2023, 54(10):407 – 415.
LU Zhixiong, WANG Yutong, WANG Lin, et al. Temperature prediction of friction pair of HMCVT wet clutch based on IGWPSO – SVM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10):407 – 415. (in Chinese)
[29] 罗琦, 茹晓雅, 姜元,等. 基于机器学习与气象灾害指标的苹果相对气象产量预测[J]. 农业机械学报, 2023, 54(9): 352 – 364.
LUO Qi, RU Xiaoya, JIANG Yuan, et al. Relative meteorological yield prediction of apples based on machine learning and meteorological hazard indicators[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9):352 – 364. (in Chinese)
[30] BHANDARI H N, RIMAL B, POKHREL N R, et al. LSTM – SDM: an integrated framework of LSTM implementation for sequential data modeling[J]. Software Impacts, 2022, 14: 100396.
[31] MOHAPATRA A, THOTA N, PRAKASAM P. Fake news detection and classification using hybrid BiLSTM and self-attention model[J]. Multimed Tools Appl., 2022, 81: 18503 – 18519.

(上接第 164 页)

[19] 魏子茹,卢延辉,王鹏宇,等. 基于 CRITIC 法的灰色关联理论在无人驾驶车辆测试评价中的应用[J]. 机械工程学报, 2021,57(12):99 – 108.
WEI Ziru, LU Yanhui, WANG Pengyu, et al. Application of grey relational theory based on CRITIC method in autonomous vehicles test and evaluation[J]. Journal of Mechanical Engineering, 2021,57(12): 99 – 108. (in Chinese)
[20] 靳亚红,吴鑫淼,甄文超,等. 基于采样点光谱信息窗口尺度优化的土壤含水率无人机多光谱遥感反演[J]. 农业机械学报,2024,55(1):316 – 327.
JIN Yahong, WU Xinmiao, ZHEN Wenchao, et al. UAV multispectral remote sensing inversion of soil moisture content based on window size optimization of spectral information at sampling points[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(1): 316 – 327. (in Chinese)
[21] 刘爽,崔修杰,吕超. 考虑多影响因素的渔船单船捕捞能力评估方法[J]. 农业工程学报,2024,40(5):90 – 98.
LIU Shuang, CUI Xiujie, LÜ Chao. Method for assessing single-vessel fishing capacity considering multiple influencing factors [J]. Transactions of the CSAE, 2024,40(5): 90 – 98. (in Chinese)
[22] 徐立鸿,黄志尊,龙伟,等. 基于 MobileViT-CBAM-BiLSTM 的开放式养殖环境鱼群摄食强度分类模型[J]. 农业机械学报,2024,55(11):147 – 153.
XU Lihong, HUANG Zhizun, LONG Wei, et al. Classification model of fish feeding intensity based on MobileViT-CBAM-BiLSTM[J]. Transactions of the Chinese Society for Agricultural Machinery ,2024,55(11):147 – 153. (in Chinese)
[23] FLUNKERT V, SALINAS D, GASTHAUS J. DeepAR: probabilistic forecasting with autoregressive recurrent networks[J]. International Journal of Forecasting, 2020, 36(3):114 – 128.
[24] 李浩,王卓健,李哲,等. 基于堆栈自编码器和 DeepAR 的航空发动机剩余寿命预测[J]. 推进技术,2022,43(11):72 – 80.
LI Hao, WANG Zhuojian, LI Zhe, et al. Prediction of remaining useful life of aero-engine based on stacked autoencoder and DeepAR[J]. Propulsion Technology, 2022,43(11): 72 – 80. (in Chinese)
[25] 何旭辉,段泉成,严磊. 基于 DeepAR 的短期风速概率预测[J]. 铁道学报,2023,45(7):152 – 160.
HE Xuhui, DUAN Quancheng, YAN Lei. Short term wind speed probability prediction model using DeepAR[J]. Journal of the China Railway Society, 2023,45(7): 152 – 160. (in Chinese)