

# AI-Based Aquaculture Recommendation System Using Machine Learning

<sup>1</sup>Naga Raju Pechetti

<sup>1</sup>M.Tech, Department of Computer Science and Engineering  
DNR College of Engineering & Technology (Autonomous)  
Bhimavaram – 534202, Andhra Pradesh, India  
nagarajucse2513@gmail.com

<sup>2</sup>Mr. B. Nandana Kumar

<sup>2</sup>Assistant Professor, Department of Computer Science and Engineering  
DNR College of Engineering & Technology (Autonomous)  
Bhimavaram – 534202, Andhra Pradesh, India  
nandan@dnrcet.org

**Abstract**— Aquaculture productivity and sustainability are strongly influenced by water-quality conditions, making continuous monitoring and timely management essential. Conventional monitoring approaches often depend on periodic measurements and manual interpretation, limiting their ability to identify rapidly changing conditions. Recent advances in machine learning provide opportunities for water-quality prediction, forecasting, anomaly identification, and intelligent aquaculture management. This survey reviews machine-learning and deep-learning approaches for aquaculture water-quality assessment, considering parameters such as pH, temperature, dissolved oxygen, turbidity, and related environmental indicators. Existing approaches including traditional machine learning, neural networks, hybrid models, time-series forecasting, and sensor-assisted monitoring are examined with respect to their capabilities and limitations. Based on the surveyed literature, an integrated AI-based aquaculture recommendation framework is presented that connects data preprocessing, water-quality prediction, anomaly detection, and management-oriented recommendations. The survey further discusses evaluation requirements, research challenges, and future directions toward intelligent, adaptive, and sustainable aquaculture systems.

**Keywords**— Aquaculture, Water Quality Prediction, Machine Learning, Deep Learning, Water Quality Forecasting, Anomaly Detection, IoT-Based Monitoring, Aquaculture Recommendation, Intelligent Aquaculture, Sustainable Aquaculture.

## I. INTRODUCTION

Aquaculture is an important component of global food production, contributing to food security, employment, economic development, and the growing demand for aquatic products. Its productivity and sustainability strongly depend on maintaining suitable water-quality conditions. Parameters such as dissolved oxygen, temperature, pH, turbidity, and other physicochemical characteristics directly affect the growth, health, and survival of cultured aquatic organisms. These parameters can change due to feeding practices, stocking density, biological activity, weather conditions, and environmental variations. Recent advances in artificial intelligence and machine learning have created opportunities for intelligent water-quality monitoring and predictive aquaculture management [1]–[3].

Conventional aquaculture water-quality assessment generally relies on manual sampling, periodic measurements, laboratory analysis, or direct observation of sensor readings. Although these methods provide useful information, they have limited capability for continuous analysis and future-condition forecasting. Machine-learning techniques can learn relationships among historical water-quality parameters and generate predictions from environmental observations [3], [12]. Deep-learning and hybrid neural-network models can further identify complex nonlinear and temporal patterns in aquaculture environments [4]–[6], [11].

Water-quality forecasting is important because aquaculture conditions can change continuously. Time-series and recurrent learning approaches have therefore been explored to predict environmental parameters and provide early warnings before critical conditions develop [5]–[7]. Sensor-assisted monitoring and UAV-based multispectral imaging have also improved environmental data collection [8], [9]. However, sensor noise, missing observations, limited datasets, environmental variability, computational requirements, and differences between facilities remain important challenges [6], [7], [10]. Prediction alone cannot provide complete intelligent aquaculture management. A practical AI-based system should integrate data acquisition, preprocessing, prediction, anomaly detection, early warning, and management recommendations. This survey therefore aims to:

1. Review machine-learning and deep-learning approaches for aquaculture water-quality assessment and prediction.
2. Examine forecasting techniques for dissolved oxygen, temperature, pH, turbidity, and related parameters.
3. Analyze sensor-assisted, UAV-based, and data-driven monitoring approaches.
4. Compare the strengths, limitations, and applicability of existing AI methods.
5. Examine anomaly detection, early-warning, and prediction-based decision support.
6. Formulate an integrated AI-based aquaculture recommendation framework.
7. Identify evaluation requirements, research challenges, and future directions for reliable and sustainable aquaculture systems.

The remainder of this paper discusses existing monitoring and prediction approaches, the proposed comprehensive framework, evaluation requirements, challenges, future research directions, and concluding observations.

## II. EXISTING METHODS FOR AI-BASED

### AQUACULTURE WATER-QUALITY MANAGEMENT

Existing aquaculture water-quality management approaches can be broadly classified into manual monitoring, sensor-based monitoring, statistical methods, machine-learning approaches, deep-learning and hybrid forecasting systems, and intelligent/remote monitoring frameworks.

#### A. Manual Water-Quality Monitoring

Manual monitoring depends on periodic sampling, laboratory analysis, and direct measurement of parameters such as pH, temperature, dissolved oxygen, and turbidity. It can provide reliable measurements and human interpretation, but frequent sampling becomes difficult for continuous aquaculture environments. Delayed measurements and increased

260 | Page

The intelligent analysis layer applies machine-learning and deep-learning techniques to identify abnormal environmental conditions and important aquaculture patterns. The prediction layer estimates future water-quality conditions, fish growth, disease risk, or other farm-related outcomes based on historical and real-time information. The risk-assessment layer combines predicted conditions to generate actionable alerts for potentially harmful situations.

Finally, the decision-support layer presents predictions, detected abnormalities, and recommended interventions to farmers through an accessible monitoring interface. The architecture is designed as a decision-support framework, assisting farmers in timely management rather than completely replacing human judgement. This integrated approach enables continuous monitoring, early identification of risks, predictive management, and more efficient aquaculture operations.

#### IV. AQUACULTURE DATA AND ENVIRONMENTAL REPRESENTATION

##### A. Aquaculture Data Acquisition and Processing

The proposed system collects aquaculture parameters such as temperature, pH, dissolved oxygen, turbidity, ammonia, nitrate, and salinity. Preprocessing handles missing, duplicate, inconsistent, abnormal data and normalizes numerical values.

The structured aquaculture representation is defined as:

$$A_i = \{W_i, F_i, E_i, H_i\} \quad (1)$$

where  $W_i$  represents water-quality parameters,  $F_i$  represents fish-related observations,  $E_i$  represents environmental conditions, and  $H_i$  represents health-related indicators.

Water-quality, fish-related, and health indicators are analyzed to assess aquaculture conditions. Normalization converts parameters with different scales and units into comparable ranges, improving joint model learning and prediction.

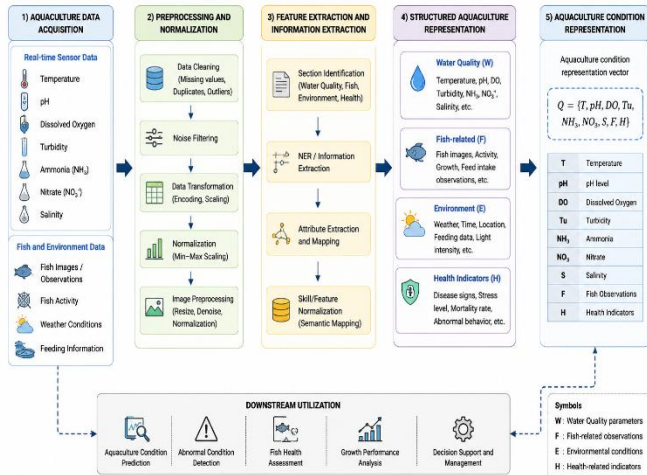


Fig. 2. Water-Quality Data Processing and Aquaculture Recommendation Generation

##### B. Aquaculture Condition Representation

The environmental condition of an aquaculture system is represented using the relevant water-quality and operational parameters. The system extracts:

- Water temperature
- pH level
- Dissolved oxygen
- Turbidity
- Ammonia level
- Nitrate level

- Salinity
- Fish-related observations
- Environmental conditions
- Health indicators

The aquaculture condition representation is defined as:

$$Q = \{T, pH, DO, Tu, NH_3, NO_3, S, F, H\} \quad (2)$$

where  $T$  represents temperature,  $pH$  represents acidity/alkalinity,  $DO$  represents dissolved oxygen,  $Tu$  represents turbidity,  $NH_3$  represents ammonia concentration,  $NO_3$  represents nitrate concentration,  $S$  represents salinity,  $F$  represents fish-related observations, and  $H$  represents health indicators.

Representing the collected aquaculture parameters using a unified structure enables the proposed system to analyze the relationship between water quality, environmental conditions, and fish health. The resulting representation can subsequently be provided to the machine-learning or deep-learning model for aquaculture condition prediction, abnormal-condition detection, or fish-health assessment, depending on the objective of the proposed system.

#### V. SEMANTIC AQUACULTURE CONDITION-MANAGEMENT MATCHING AND FARM SCORING

A central component of the proposed aquaculture system is semantic condition-management matching. Unlike simple threshold-based analysis, the proposed approach identifies relationships between the observed aquaculture conditions and the recommended farming requirements. The system considers parameters such as water quality, temperature, pH, dissolved oxygen, turbidity, ammonia, feeding conditions, and other available farm parameters to determine the suitability of the current farming environment.

Let the aquaculture condition embedding and recommended management embedding be:

$$V_C = \text{Embedding}(C_i)$$

$$V_R = \text{Embedding}(R)$$

Their semantic similarity can be calculated using cosine similarity:

$$S_{sem} = \frac{V_C \cdot V_R}{\|V_C\| \|V_R\|} \quad (3)$$

A higher value indicates stronger compatibility between the observed aquaculture conditions and the recommended management conditions.

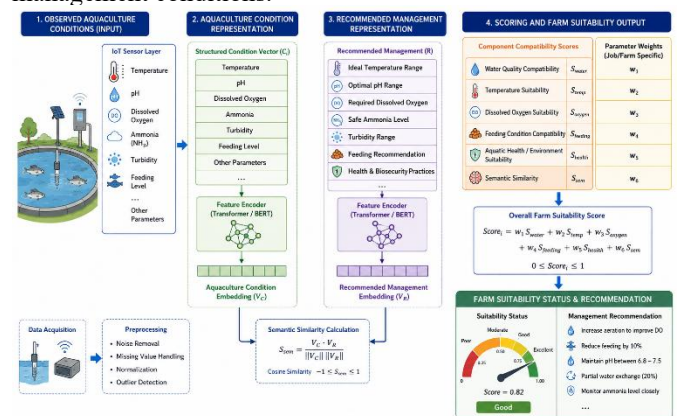


Fig. 3. Semantic Aquaculture Condition-Management Matching and Farm Suitability Scoring

Semantic similarity alone, however, should not determine the final aquaculture decision. A farm condition with high overall similarity may still contain a critical parameter, such as insufficient dissolved oxygen or excessive ammonia, that can negatively affect aquatic organisms. Therefore, the proposed system combines semantic similarity with explicit aquaculture-specific factors.

The overall farm suitability score can be defined as:

$$\text{Score}_i = w_1 \text{Swater} + w_2 \text{Stemp} + w_3 \text{Soxygen} + w_4 \text{Sfeeding} + w_5 \text{Shealth} + w_6 \text{Ssem}(4)$$

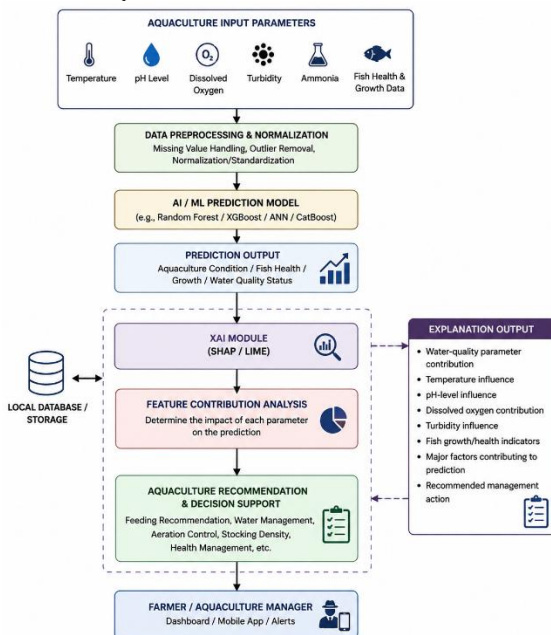
where:

- Swater = water-quality compatibility,
- Stemp = temperature suitability,
- Soxygen = dissolved-oxygen suitability,
- Sfeeding = feeding-condition compatibility,
- Shealth = aquatic-health/environmental suitability,
- Ssem = semantic similarity between observed and recommended conditions,
- $w_1, \dots, w_6$  = parameter-specific weights.

This weighted formulation allows the aquaculture management system to assign different importance to different environmental and operational conditions. For example, a fish-growth monitoring application may assign greater importance to temperature, dissolved oxygen, pH, and feeding conditions, whereas a water-quality management application may emphasize pH, ammonia, turbidity, dissolved oxygen, and temperature.

The aquaculture conditions are subsequently evaluated according to their final scores to generate a farm suitability status and corresponding management recommendation, enabling timely intervention and improved aquaculture management.

## VI. EXPLAINABLE AI AND DECISION-AWARE AQUACULTURE MANAGEMENT



**Fig. 4. Explainable AI Framework for Aquaculture Prediction and Decision Support**

AI-based aquaculture systems should provide more than a prediction or numerical output. Farmers and aquaculture

managers need to understand why the system generated a particular prediction or recommendation. Explainable AI (XAI) techniques such as SHAP and LIME can provide feature-level explanations of model predictions.

For each prediction, the proposed system can generate an explanation containing:

1. Water-quality parameter contribution
2. Temperature influence
3. pH-level influence
4. Dissolved oxygen contribution
5. Turbidity or water-condition influence
6. Fish growth or health-related indicators
7. Major factors contributing to the final prediction
8. Recommended aquaculture management action

For example:

### **Aquaculture Condition — 91% Suitable**

- Water quality: highly suitable.
- Temperature: within the optimal range.
- pH: suitable.
- Dissolved oxygen: adequate.
- Turbidity: acceptable.
- Fish condition: healthy.
- Major influencing factor: stable water-quality parameters.
- Recommendation: Maintain current pond conditions and continue regular monitoring.

Such explanations help farmers understand whether the AI prediction is supported by meaningful aquaculture parameters rather than relying only on the final prediction value.

### **A. Explainability and Decision Monitoring**

Aquaculture prediction models can be influenced by multiple environmental and biological parameters. Therefore, the system should continuously monitor the contribution of each input feature to the prediction. Explainability helps identify which parameters have the greatest influence on fish health, growth, water quality, or other target outcomes considered by the proposed system.

The proposed framework therefore integrates explainability as a decision-support component. The system should report both: Performance = How accurately does the system predict the required aquaculture condition?

Explainability = Can the farmer understand which aquaculture parameters caused the prediction?

This dual objective is important because a highly accurate model that provides no understandable reasoning may be difficult to trust or use in practical aquaculture environments.

## VII. ADAPTIVE AQUACULTURE INTELLIGENCE AND HUMAN-IN-THE-LOOP DECISION SUPPORT

Aquaculture conditions are not static. Water temperature, pH, dissolved oxygen, turbidity, feeding requirements, fish growth, and other environmental factors continuously change during the culture cycle. An aquaculture prediction model trained using fixed environmental conditions may therefore become less effective when pond conditions change.

### **Model Update**

Adaptive mechanisms can update aquaculture parameter ranges, environmental thresholds, feature importance, prediction knowledge, and management recommendations while maintaining model version control.

The system can periodically incorporate newly collected pond data to improve its understanding of changing aquaculture

conditions. However, adaptive AI should not automatically modify critical management recommendations without validation. Model updates should first be evaluated using historical and newly collected aquaculture datasets before being deployed.

#### Human-in-the-Loop Workflow

The final aquaculture decision-support workflow is:

Data Collection → AI Prediction → Condition Assessment → Explanation → Farmer Verification → Management Recommendation → Final Farmer Decision

The farmer or aquaculture expert remains responsible for the final decision, particularly when environmental conditions are unusual or when additional field information is required that may not be available in the collected data.

This human-in-the-loop design allows farmers to combine AI-based predictions with practical aquaculture knowledge and real-time pond observations. Farmer feedback can also be captured and used for future system improvement, while preventing the AI system from making completely autonomous decisions without human supervision.

### VIII. EVALUATION FRAMEWORK

A comprehensive evaluation framework is adopted to assess the proposed recruitment system across screening, ranking, semantic matching, explainability, fairness, efficiency, and robustness. Rather than relying on a single performance measure, the evaluation considers multiple complementary dimensions to determine the practical effectiveness of the proposed framework.

#### A. Screening Performance

The screening module is evaluated using accuracy, precision, recall, and F1-score. Accuracy measures overall classification correctness, while precision indicates the proportion of selected candidates who are relevant. Recall measures the ability to identify qualified candidates, and F1-score provides a balanced assessment of precision and recall. These metrics are important because both incorrect selections and missed qualified candidates can reduce recruitment effectiveness.

#### B. Ranking Performance

Since applicant ranking is the primary objective, the system is evaluated using Precision@K, Recall@K, Mean Reciprocal Rank (MRR), and Normalized Discounted Cumulative Gain (NDCG). These metrics determine whether suitable candidates are positioned within the most important top-ranking positions and whether highly relevant applicants are prioritized appropriately.

#### C. Semantic Matching

The semantic matching component is compared with a conventional keyword-based baseline. The evaluation examines whether contextual representations can identify relevant candidates despite differences in terminology, wording, skills, experience descriptions, and professional expressions. This assessment determines the effectiveness of contextual candidate-job alignment.

#### D. Explainability

The generated candidate explanations are evaluated based on faithfulness, consistency, completeness, and human understandability. The explanation should accurately represent the factors influencing the ranking decision and enable recruiters to verify skills, experience, education, projects, certifications, and missing requirements.

#### E. Fairness

Fairness evaluation examines potential differences in ranking and selection outcomes across controlled demographic groups. Controlled testing can determine whether modifying irrelevant demographic information produces unjustified changes in candidate recommendations. Thus, both predictive performance and responsible decision-making are considered.

#### F. Efficiency and Robustness

Computational evaluation considers resume processing time, ranking latency, memory consumption, throughput, and scalability. Robustness is evaluated through cross-domain testing, different resume formats, terminology variations, incomplete information, and diverse job categories. This ensures that the proposed system maintains reliable performance under practical recruitment conditions.

TABLE II EVALUATION DIMENSIONS OF THE PROPOSED RECRUITMENT FRAMEWORK

| Dimension         | Evaluation Metrics / Criteria                                        | Purpose                                       |
|-------------------|----------------------------------------------------------------------|-----------------------------------------------|
| Screening         | Accuracy, Precision, Recall, F1-Score                                | Evaluate candidate identification performance |
| Ranking           | Precision@K, Recall@K, MRR, NDCG                                     | Evaluate applicant ranking quality            |
| Semantic Matching | Semantic similarity, Keyword Baseline Comparison                     | Assess candidate-job alignment                |
| Explainability    | Faithfulness, Consistency, Completeness, Human Understandability     | Evaluate decision transparency                |
| Fairness          | Group disparity, Controlled demographic evaluation                   | Identify potential recruitment bias           |
| Efficiency        | Processing Time, Ranking Latency, Memory, Throughput                 | Assess computational efficiency               |
| Robustness        | Cross-domain Testing, Resume Format Variation, Terminology Variation | Evaluate generalization capability            |

### IX. RESEARCH CHALLENGES AND FUTURE ROADMAP

Although AI-powered recruitment can improve candidate screening and job matching, several challenges remain in developing a reliable and practical recruitment system.

#### A. Dataset Availability

Recruitment datasets may contain limited candidate profiles, incomplete professional information, and variations in resume formats. Developing sufficiently diverse and representative datasets is therefore important for reliable candidate-job matching.

#### B. Candidate-Job Matching

Candidates may describe similar skills, qualifications, and experience using different terminology. The proposed system must therefore handle variations in skills, job roles, experience descriptions, and professional terminology to achieve reliable matching.

#### C. Ranking Reliability

Candidate ranking should consider multiple job-relevant factors rather than depending on a single similarity score. Future improvements should focus on maintaining consistent rankings across different job descriptions and candidate profiles.

#### D. Explainability

AI-generated candidate rankings should provide understandable reasons for recommendations. Future systems should improve the transparency of skill matching, experience

relevance, educational compatibility, project relevance, and certification matching.

#### **E. Fairness**

Recruitment systems should minimize inappropriate influence from demographic or unrelated candidate attributes. Fairness-aware evaluation is therefore required to ensure that candidate ranking primarily reflects job-relevant qualifications.

#### **F. Skill and Job Evolution**

New technologies, tools, certifications, and professional roles continuously emerge. Future recruitment systems should therefore support updated skill representations and evolving job requirements without reducing previously learned capabilities.

#### **G. Human-AI Interaction**

AI recommendations should assist recruiters rather than completely replace human judgement. Future systems should provide effective recruiter feedback mechanisms and decision-support interfaces for reviewing and validating candidate recommendations.

#### **H. Privacy and Security**

Resume information contains personal and professional details. Secure data management, controlled access, anonymization, and privacy-preserving processing are therefore important for practical recruitment deployment.

The research roadmap can therefore be organized into three stages:

**Short Term:** Improve resume parsing, semantic candidate-job matching, multi-factor ranking, explainability, and fairness evaluation.

**Medium Term:** Develop adaptive skill representations, improved job-specific ranking, diverse recruitment datasets, and robust cross-domain evaluation.

**Long Term:** Develop reliable human-centered recruitment intelligence that combines semantic matching, explainable ranking, fairness-aware evaluation, adaptive learning, and human-in-the-loop decision support.

#### **X. CONCLUSION**

This paper presented an AI-powered framework for intelligent resume screening and job applicant ranking. The framework addresses the limitations of conventional keyword-based recruitment systems by incorporating resume processing, job-description analysis, semantic candidate-job matching, multi-factor suitability scoring, applicant ranking, explainability, fairness monitoring, and human-in-the-loop decision support.

The proposed approach evaluates candidates using job-relevant skills, experience, education, projects, certifications, and other relevant qualifications. Semantic matching enables the system to identify contextual relationships between candidate profiles and job requirements beyond exact keyword matching. The ranking component further prioritizes candidates according to their overall suitability for the specified position.

Explainability and fairness are incorporated to improve the transparency and reliability of recruitment recommendations. Recruiters can examine the major factors contributing to candidate rankings, while fairness evaluation helps identify potentially inappropriate influences on recruitment outcomes. The human-in-the-loop design ensures that AI-generated recommendations support recruiter decision-making rather than completely replacing human judgement.

Future research should focus on diverse recruitment datasets, adaptive skill representations, robust ranking evaluation, faithful explainability, fairness-aware learning, privacy-

preserving processing, and reliable AI-based recruitment decision support.

#### **REFERENCES**

- [1] S. Zhao, S. Zhang, J. Liu, H. Wang, J. Zhu, D. Li, and R. Zhao, "Application of machine learning in intelligent fish aquaculture: A review," *Aquaculture*, vol. 540, Art. no. 736724, 2021, doi: 10.1016/j.aquaculture.2021.736724.
- [2] J. Bi, Y. Lin, Q. Dong, H. Yuan, and M. C. Zhou, "Large-scale water quality prediction with integrated deep neural network," *Information Sciences*, vol. 571, pp. 191–205, 2021, doi: 10.1016/j.ins.2021.04.057.
- [3] "Predicting aquaculture water quality using machine learning approaches," *Water*, vol. 14, no. 18, Art. no. 2836, 2022, doi: 10.3390/w14182836.
- [4] J. Yang, L. Jia, Z. Guo, Y. Shen, X. Li, Z. Mou, K. Yu, and J. C.-W. Lin, "Prediction and control of water quality in recirculating aquaculture system based on hybrid neural network," *Engineering Applications of Artificial Intelligence*, vol. 121, Art. no. 106002, 2023, doi: 10.1016/j.engappai.2023.106002.
- [5] E. Eze, S. Kirby, J. Attridge, and T. Ajmal, "Aquaculture 4.0: Hybrid neural network multivariate water quality parameters forecasting model," *Scientific Reports*, vol. 13, Art. no. 16129, 2023, doi: 10.1038/s41598-023-41602-7.
- [6] X. Zhou, J. Wang, Y. Liu, and Q. Duan, "Deep learning with PID residual elimination network for time-series prediction of water quality in aquaculture industry," *Computers and Electronics in Agriculture*, vol. 212, Art. no. 108125, 2023, doi: 10.1016/j.compag.2023.108125.
- [7] R. Jongjaraunsuk, W. Taparhudee, and P. Suwannasing, "Comparison of water quality prediction for red tilapia aquaculture in an outdoor recirculation system using deep learning and a hybrid model," *Water*, vol. 16, no. 6, Art. no. 907, 2024, doi: 10.3390/w16060907.
- [8] "Monitoring water quality parameters of freshwater aquaculture ponds using UAV-based multispectral images," *Ecological Indicators*, vol. 167, Art. no. 112644, 2024, doi: 10.1016/j.ecolind.2024.112644.
- [9] "Estimating water quality parameters of freshwater aquaculture ponds using UAV-based multispectral images," *Agricultural Water Management*, vol. 304, Art. no. 109088, 2024, doi: 10.1016/j.agwat.2024.109088.
- [10] "Integrating external stressors in supervised machine learning algorithm achieves high accuracy to predict multi-species biological integrity index of aquaculture wastewater," *Journal of Hazardous Materials*, vol. 480, Art. no. 136366, 2024, doi: 10.1016/j.jhazmat.2024.136366.
- [11] Q. Ren, X. Wang, W. Li, Y. Wei, and D. An, "Research of dissolved oxygen prediction in recirculating aquaculture systems based on deep belief network," *Aquacultural Engineering*, vol. 90, Art. no. 102085, 2020, doi: 10.1016/j.aquaeng.2020.102085.
- [12] Y. Khan and C. S. See, "Predicting and analyzing water quality using machine learning: A comprehensive model," in *Proc. IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, 2016, pp. 1–6.
- [13] D. Dheda and L. Cheng, "A multivariate water quality parameter prediction model using recurrent neural network," *arXiv preprint arXiv:2003.11492*, 2020.