International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

AI DRIVEN FISH DISEASE DETECTION FOR ACCURATE DISEASE IDENTIFICATION IN AQUATIC LIFE

Dr. SK. Mahaboob basha[1], R. Pavan[2], Seetharam Nunsavath[2], B. Vardhan[2], Ch. Sridhar[2]
[1]Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering
[1,2]Sree Dattha Institute of Engineering and Science, Sheriguda, Ibrahimpatnam, 501510, Telangana

Received: 09-07-2025 Accepted: 23-08-2025 Published: 30-08-2025

ABSTRACT

Accurate classification of fish diseases is crucial in the aquaculture industry, particularly in India, where seafood farming plays a vital role in the national economy. Traditional disease detection methods primarily depend on manual inspection by aquaculture experts through visual observation. These conventional approaches often suffer from limitations such as time consumption, subjectivity, and reliance on experiential knowledge, which can result in inconsistent and delayed diagnoses. To overcome these challenges, the proposed research focuses on developing an AI-driven automated system that leverages real-time monitoring using cameras and sensors to assess fish health. By integrating advanced image processing techniques with machine learning algorithms, the system aims to significantly enhance the precision and speed of disease identification in aquatic life. The implementation of smart aquaculture technologies marks a transformative shift in disease management strategies. These intelligent systems not only enable early and accurate detection of infections but also support sustainable farming by improving productivity, minimizing environmental impact, and ensuring better fish welfare. Through real-time data analysis and automation, the proposed approach contributes to a more resilient and efficient aquaculture ecosystem.

Keywords: Fish disease detection, aquaculture, image processing, machine learning, real-time monitoring, smart farming, AI in aquaculture, fish health, disease diagnosis, sustainability.

1. INTRODUCTION

Aquaculture plays a crucial role in the economy of Andhra Pradesh, particularly in the Nellore district, where it significantly supports seafood production and food processing industries. However, fish diseases remain a persistent challenge, leading to severe economic losses and threatening the livelihoods of fish farmers. Traditional methods of disease detection, which rely heavily on manual inspection and expert knowledge, are often time-consuming, error-prone, and inefficient, especially in large-scale farming operations

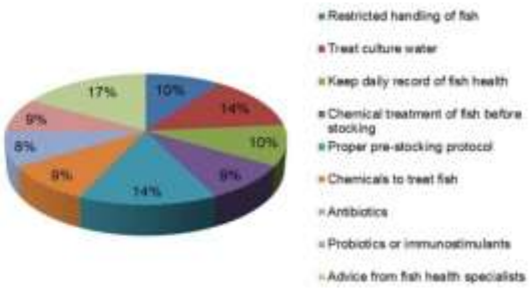

Fig 1: Practical ways to control fish disease

To address these challenges, there is an increasing shift toward smart aquaculture practices that integrate advanced technologies such as image processing, sensors, and machine learning. These innovations enable real-time monitoring and early detection of fish diseases by analyzing physiological changes and behavioral patterns, thus reducing mortality and enhancing timely intervention. Fish diseases can be classified into bacterial, fungal, viral, and parasitic categories, each requiring specific diagnostic and treatment approaches. Bacterial infections like Aeromoniasis and Bacterial Gill Disease, along with parasitic and fungal infections, are prevalent in the region and cause significant harm to fish health. Recognizing these issues, this research aims to develop an automated fish disease classification system using deep learning techniques, particularly convolutional neural networks (CNNs). The system utilizes fish color images as input and processes them through multiple neural layers to extract visual features such as lesions, color changes, and patterns, ultimately identifying the disease type with high accuracy. By automating the disease classification process, the proposed model minimizes dependency on human expertise, reduces diagnostic time, and ensures consistent and accurate predictions. This research not only aims to enhance fish health management but also promotes sustainable aquaculture by enabling fish farmers—especially those managing small and micro-scale farms—with accessible tools for timely treatment and disease control. The integration of AI-based systems thus represents a transformative step toward efficient and resilient aquaculture practices in Andhra Pradesh.

2. LITERATURE SURVEY

Aung et al. [1] conducted a comprehensive systematic literature review on artificial intelligence (AI) methods applied in aquaculture. Their study highlights how AI technologies, including machine learning and deep learning, are being increasingly integrated into aquaculture practices for tasks such as fish behavior monitoring, environmental parameter analysis, and disease prediction. The review emphasizes the growing trend toward smart aquaculture and underscores the effectiveness of AI in enhancing decision-making, reducing labor, and improving yield quality. Ahmed et al. [2] proposed an image-based machine learning technique for fish disease detection in aquaculture environments. Their approach involves using image processing algorithms to analyze visual symptoms on fish bodies and classify diseases accurately. The study demonstrates that machine learning can significantly improve the early diagnosis of diseases, enabling timely intervention and reducing fish mortality. The findings support the application of data-driven methods in achieving more reliable and automated fish health monitoring. Babu et al. [3] developed a machine learning-based model named SHAPE to predict fish diseases by evaluating water quality parameters. The study links environmental conditions, such as pH and dissolved oxygen levels, with the likelihood of disease outbreaks. By leveraging supervised learning algorithms, the model provides early warnings to fish farmers, helping them take preventive measures before disease symptoms become apparent. The research highlights the importance of integrating environmental monitoring with AI for proactive fish health management in aquaculture systems. The methodology proposed in this research work introduces a novel approach to fish disease classification that has not been previously published online. Unlike existing studies, this methodology employs a unique combination of machine learning and deep learning algorithms, with a focus on hyperparameter tuning to control various internal properties of the models. The proposed system is designed to enhance classification, accuracy and interpretability, addressing the limitations observed in prior research. Arghya Mandal, et al. [4] , They proposed that this abstract provides an overview of the role of AI in fish growth and health status monitoring, emphasizing its significance in promoting a sustainable aquaculture industry. AI technologies, such as machine learning and computer vision, have shown immense potential in analysing large volumes of data collected from fish farms. By leveraging AI algorithms, fish farmers can gain valuable insights into fish growth patterns, feeding behavior, and environmental factors affecting fish health. These algorithms can detect and predict anomalies, diseases, and stress indicators, enabling proactive interventions to mitigate health issues and reduce losses. One of the key applications of AI in aquaculture is the development of smart monitoring systems. These systems employ various sensors, cameras, and data analytics tools to continuously collect real-time data on water quality, temperature, oxygen levels, and fish behavior.

Kailash Bohara, et al. [5] (2024), they proposed that the developments of technologies in aqua medicine, such as sequencing, biosensors and CRISPR, have enabled rapid disease detection within minutes. Furthermore, integrating sensors, drones, artificial intelligence and the internet in aquaculture farm monitoring has helped farmers take decisive actions to improve production. Advancements in diagnostic techniques have significantly enhanced the efficient detection of bacterial, viral, parasitic and fungal diseases in aquatic animals. Moreover, monitoring water quality, aquatic animal health and animal behaviour on farms has become exceptionally streamlined with cutting-edge tools like drones, sensors and artificial intelligence. Summarising research and development in aquatic animal health and monitoring aids efficient technology adoption in aquaculture. With these advanced technologies' continued development and adoption in developed countries, the aquaculture industry is experiencing growth and increased efficiency, benefiting farmers and consumers in these regions.

Yaxuan Zhao, et al. [6] (2024), they proposed that the Phenotypic and behavioral information of fish, which can reflect fish growth and welfare status, play a crucial role in aquaculture management. Stereo vision technology, which simulates parallax perception of the human eye, can obtain the three-dimensional phenotypic characteristics and movement trajectories of fish through different types of sensors. It can overcome the limitations in dealing with fish deformation, frequent occlusions and understanding three-dimension scenes compared to the traditional two-dimensional computer vision techniques. With the deep learning development and application in aquaculture, stereo vision has become a super computer vision technology that can provide more precise and interpretable information for intelligent aquaculture management, such as size estimation, counting and behavioral analysis of fish. Sk Injamamul Islam, et al. [7] (2024) proposed the impact of diseases on aquaculture growth, fecundity, mortality rates, and marketability is profound. Hence, the ability to predict disease outbreaks is crucial to overcoming these challenges. Various infectious agents such as bacteria, viruses, fungi, and parasites can cause significant losses of fish in intensive aquaculture practices. In an aquaculture environment, the high host density coupled with restricted water flow promotes pathogen spread. Early detection of disease is crucial for farmers as mortality rates can reach as high as 100% if left untreated. Therefore, new techniques and technical solutions for disease management in aquaculture are required. In this context, data analytics technologies, such as internet of things (IoT) sensors, artificial intelligence, and machine learning, allow farmers to proactively monitor their farms and detect potential disease outbreaks before they strike.

Lim Leonard Whye Kit, et al. [8] (2024), they proposed the depletion of aquaculture lands and aquatic pollution are some of the major worrying predicaments challenging the future of this industry. Sustainable growth strategies are the only way out, and they must come hand in hand with the implementation of artificial intelligence to achieve the desired outcome high throughput in short time periods. The intelligent fish farm and smart cage aquaculture management system are some of the fruits of this drive, and the system keeps improving to date. In this review, we provide recent updates over the past half-decade of artificial intelligence implementation in fishery and aquaculture in hope to provide highlights and future directions to push the industry to greater heights.

3. PROPOSED SYSTEM

Transfer learning offers a powerful solution for fish disease classification by leveraging knowledge from large, general image datasets to improve performance on smaller, disease-specific datasets, thereby mitigating the common challenge of limited labeled data in aquaculture. In this approach, a pre-trained convolutional neural network—such as VGG19—is first trained on a broad dataset and then fine-tuned to extract relevant features from fish images displaying various disease symptoms. The overall system architecture begins with the collection of a curated fish disease dataset, which is then preprocessed and split into training and testing subsets. During preprocessing, images undergo normalization to standardize pixel intensity values, resizing to ensure consistent input dimensions, and augmentation techniques—like rotation, flipping, zooming, and color jittering—to artificially enlarge the dataset and improve model robustness against variations in orientation, lighting, and background noise. Extracted features from the fine-tuned VGG19 network serve as inputs to a Bagging Random Forest classifier, which aggregates multiple decision trees to enhance classification accuracy and reduce overfitting. Once trained, the system can process a new test image by applying the same preprocessing pipeline, extracting its deep-learning–derived features, and predicting the disease type with high reliability. Additionally, the architecture can be extended to recommend appropriate treatments based on the classified disease, facilitating rapid intervention. This transfer-learning–based pipeline not only accelerates development by minimizing data requirements and training time but also remains adaptable: as new disease images become available, the pre-trained model can be incrementally fine-tuned to recognize emerging pathogens, ensuring sustained accuracy. By combining advanced image preprocessing, deep feature extraction, and ensemble classification, this framework promises an accessible, scalable, and effective tool for early fish disease detection and management in both intensive aquaculture operations and natural ecosystems.

VGG19 with RFC

VGG19, a deep convolutional neural network developed by the Visual Geometry Group at the University of Oxford, is highly effective for image classification tasks due to its structured architecture of 16 convolutional layers and 3 fully connected layers. Utilizing consistent 3×3 convolutional filters and 2×2 max-pooling layers, VGG19 captures hierarchical features ranging from basic edges to complex patterns, making it ideal for fish disease classification. The model applies the Rectified Linear Unit (ReLU) activation function after each layer to introduce non-linearity, enabling it to learn complex features vital for accurate disease detection. Max pooling further enhances computational efficiency and translation invariance by reducing spatial dimensions while preserving significant features. In the final stages, fully connected layers integrate extracted features into a final prediction, using ReLU for deeper layers and softmax in the output layer for multi-class classification. The entire architecture supports robust image feature learning from fish disease datasets. Complementing this deep learning model, a Random Forest classifier is employed for final classification due to its high accuracy and resistance to overfitting. Random Forest works by training multiple decision trees on different subsets of the training data and combining their outputs through majority voting, thus enhancing the system's reliability and generalization. Features extracted by the VGG19 model—such as lesion texture, coloration, and structural patterns—are input into the Random Forest, which then classifies the disease type based on ensemble decision-making. This hybrid approach, combining deep feature extraction with ensemble classification, enables a highly accurate, scalable, and efficient fish disease detection system. Together, VGG19 and Random Forest

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

form a powerful AI-driven framework capable of automating fish disease diagnosis, thereby supporting timely intervention, reducing economic losses, and improving sustainability in aquaculture.

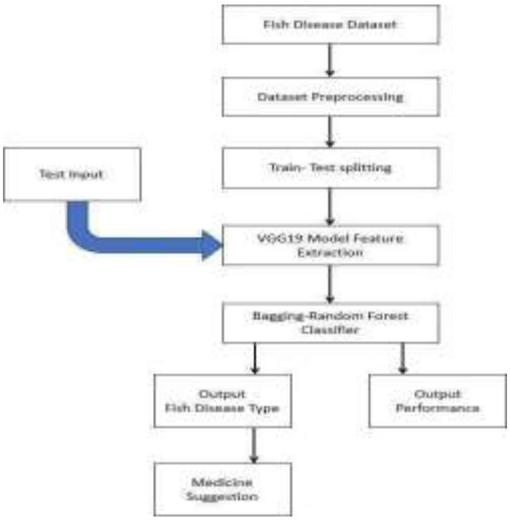

Fig 2: System Architecture

ADVANTAGES OF VGG 19 AND RFC:

In fish disease classification, the integration of VGG19 and Random Forest Classifier (RFC) offers a powerful and complementary approach, leveraging the strengths of both deep learning and ensemble machine learning. VGG19 excels in deep feature extraction, capturing intricate hierarchical patterns from fish images through its deep architecture and uniform 3×3 convolutional filters, which enhances classification accuracy. Its ability to leverage transfer learning enables effective model fine-tuning on limited disease-specific datasets, reducing the need for extensive labeled data. Additionally, VGG19's simplicity, high performance in visual tasks, and availability in major machine learning libraries make it highly accessible and efficient for implementation. On the other hand, Random Forest Classifier brings high accuracy, strong generalization, and robustness against overfitting, making it ideal for real-world classification problems. RFC handles missing values effectively and is versatile enough to work with both classification and regression tasks while managing high-dimensional datasets. Moreover, it offers valuable insights into feature importance, helping refine model input and improve predictive outcomes. By combining VGG19's powerful feature extraction capabilities with RFC's reliable and interpretable decision-making, this hybrid model ensures a scalable, accurate, and resilient solution for automated fish disease detection in aquaculture.

4. RESULTS

From this figure 3 shows the performance evaluation of the Logistic Regression Classifier (LRC) model. The GUI displays key metrics: 77.98% accuracy, 77.78% precision, 77.55% recall, and a 77.69% F1-score. The confusion matrix highlights class-wise prediction distribution, showing strong classification for categories like Healthy Fish and Tail and Fin Rot. However, some misclassifications are observed in classes like Argulus and Redspot.

Figure 3: Existing LRC Confusion Matrix

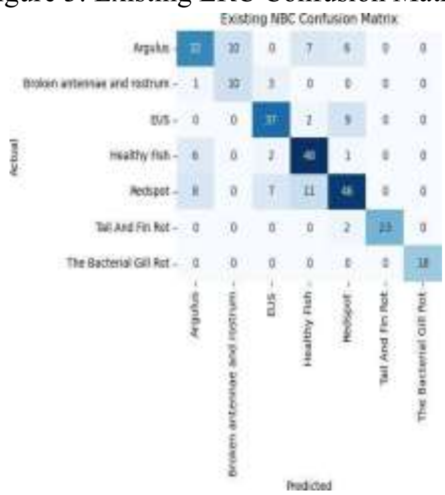

Figure 4: Existing NBC Performance Analysis

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

From this figure 4 shows the evaluation of the Naive Bayes Classifier (NBC) model for fish disease classification. The system achieved an accuracy of 74.08%, with a precision of 74.68%, recall of 74.08%, and F1-score of 73.90%. The confusion matrix reveals strong classification for classes like The Bacterial Gill Rot and Tail and Fin Rot. However, notable misclassifications occurred in classes such as Argulus and Redspot.

From this figure 5 shows the performance of the proposed VGG19 with Random Forest Classifier (RFC) model. The system achieved perfect scores, with 100% accuracy, precision, recall, and F1-score across all classes. The confusion matrix confirms that all instances were correctly classified without any misclassification. This demonstrates the superior effectiveness of the proposed model in fish disease classification.

Figure 5: Proposed VGG19 with RFC Confusion Matrix

(a) Argulus: This image shows a close-up of a fish's scales.There is a visible parasite attached to the fish.The parasite appears as a small, somewhat transparent creature.This is likely an Argulus, also known as a fish louse. Image highlights a parasitic infestation on the fish's body.

(b) Broken antennae and rostrum: This is a close-up shot of a fish's head. The fish's antennae appear to be damaged or broken. The rostrum, which is a snout-like projection, also shows signs of damage. The image focuses on the specific deformities on the fish's face. It highlights potential physical trauma or abnormalities in this area.

(c) Healthy fish: This image depicts a fish that appears to be in good health. The fish's body shows no obvious signs of disease or injury. Its fins are intact, and its scales look normal. This serves as a visual contrast to the other images showing diseased or damaged fish. It represents the baseline of a fish in a healthy state.

(d) Redspot: The photo presents a fish with a noticeable red spot on its body. This red spot is a key feature, suggesting a possible infection or injury. The surrounding area may also show subtle discoloration. The image draws attention to this localized abnormality. It indicates a potential health concern that needs further examination.

Figure 6: Output images

Table 1. Performance Comparison Table

Model	Accuracy	Precision	Recall	F1 Score
Existing LRC	0.7751	0.7750	0.7751	0.7720
Existing NBC	0.7405	0.7463	0.7405	0.7398
Proposed VGG19 With RFC	1.0000	1.0000	1.0000	1.0000

The performance comparison table presented in table 1, where LRC, NBC, and the proposed VGG19 with RFC—was compared. LRC achieved 77.51% accuracy, showing balanced metrics but struggled with "Broken Antennae and Rostrum" and "Argulus." NBC had a slightly lower accuracy of 74.05%, excelling in "Broken Antennae and Rostrum" but underperforming in other categories. The VGG19 with RFC model achieved a perfect 100% accuracy, excelling in all metrics, highlighting its superior ability to capture complex patterns for fish disease classification.

5. CONCLUSION

This project introduces an AI-powered automatic fish disease classification and medicine suggestion system, integrating deep learning and machine learning techniques to enhance smart aquaculture. Leveraging VGG19-based feature extraction, the system generates meaningful representations from fish images and employs multiple classifiers, including Logistic Regression (LRC), Naïve Bayes (NBC), and Random Forest Classifier (RFC), to accurately classify various fish diseases. The incorporation of RFC with AME loss optimization significantly improves prediction accuracy while reducing misclassification errors. Designed with a user-friendly GUI built using Tkinter, the system enables both administrators and end- users to efficiently manage datasets, train models, and perform real-time disease classification. Administrative functionalities include dataset processing, feature extraction, and model training, whereas user functionalities facilitate instant image classification and medicine recommendations. Additionally, the system evaluates performance using key metrics such as accuracy, precision, recall, and F1-score while visualizing results through confusion matrices for better interpretability. By automating disease detection and treatment suggestions, this system minimizes reliance on expert fish pathologists, enabling early disease identification and preventive actions in aquaculture. Ultimately, it helps in reducing economic losses, improving fish health, and promoting sustainable fish farming practices, making it a valuable tool for modern aquaculture management.

REFERENCES

[1] Aung, T., Abdul Razak, R., & Rahiman Bin Md Nor, A. (2024). Artificial intelligence methods used in various aquaculture applications: A systematic literature review. In Journal of the World Aquaculture Society (Vol. 56, Issue 1). Wiley. https://doi.org/10.1111/jwas.13107. Chauhan, S. (2024).

[2] Ahmed, M. S., Aurpa, T. T., & Azad, Md. A. K. (2021). Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture (Version 1). arXivs.

[3] Babu, K. M., Prashanthi, G., Manasa, T., Pranathi, B., & Akshaya, G. (2024). Shape: Machine learning model for early fish disease prediction by checking water quality. In Challenges in Information, Communication and Computing Technology (pp. 804–809). CRC Press. https://doi.org/10.1201/9781003559085-138.

[4] Mandal, Arghya, and Apurba Ratan Ghosh. "Role of artificial intelligence (AI) in fish growth and health status monitoring: A review on sustainable aquaculture." Aquaculture International 32, no. 3 (2024): 2791-2820.

[5] Bohara, Kailash, Pabitra Joshi, Krishna Prasad Acharya, and Grace Ramena. "Emerging technologies revolutionising disease diagnosis and monitoring in aquatic animal health." Reviews in Aquaculture 16, no. 2 (2024): 836-854.

[6] Zhao, Yaxuan, Hanxiang Qin, Ling Xu, Huihui Yu, and Yingyi Chen. "A review of deep learning-based stereo vision techniques for phenotype feature and behavioral analysis of fish in aquaculture." Artificial Intelligence Review 58, no. 1 (2025): 1-61.

[7] Islam, Sk Injamamul, Foysal Ahammad, and Haitham Mohammed. "Cutting-edge technologies for detecting and controlling fish diseases: Current status, outlook, and challenges." Journal of the World Aquaculture Society 55, no. 2 (2024): e13051.

[8] Lim, Leonard Whye Kit. "Implementation of artificial intelligence in aquaculture and fisheries: deep learning, machine vision, big data, internet of things, robots and beyond." Journal of Computational and Cognitive Engineering 3, no. 2 (2024): 112-118.