IJITCE

International Journal of
Information Technology & Computer Engineering

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

Uradi Rakesh [1] Sirnapally Ankith [2], Girmi Sai Kumar [3], Mrs. D. Manasa [4]

[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

[4]Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India. 500100

[4]dmanasacse@smec.ac.in

Abstract:

Fish diseases are a significant concern in aquaculture, causing economic losses and threatening biodiversity. Early and accurate detection is crucial to prevent the spread of these diseases and maintain the health of aquatic life. Traditional methods of disease detection relied on visual inspections and expert knowledge, which were time-consuming and often inaccurate. These systems, though helpful, lacked scalability and efficiency. AI-driven systems have revolutionized fish disease detection by leveraging machine learning and deep learning algorithms. These systems use large datasets and image recognition techniques to identify diseases based on visual symptoms, improving the accuracy and speed of diagnosis. The history of AI in this field began with basic image processing and evolved to more advanced models capable of accurate disease identification. Prior to AI, disease detection relied heavily on manual methods like bacterial culture tests and microscopic analysis, which were labor-intensive and slow. These methods made it difficult to diagnose diseases in a timely manner, leading to potential spread and economic loss. The introduction of AI has automated this process, enabling real-time detection and accurate diagnosis. The motivation for developing an AI-driven fish disease detection system lies in the need for efficient and scalable solutions in the aquaculture industry. The rising global demand for seafood, coupled with the challenges of maintaining fish health, necessitates technological advancements to ensure rapid disease management. AI offers an opportunity to optimize disease detection, making it faster, more accurate, and accessible. Traditional systems faced challenges such as subjectivity in diagnoses, slow diagnostic processes, and high reliance on human expertise. These issues increased costs and delayed interventions. The proposed AI-based system addresses these problems by using machine learning to automate the detection process, providing timely and accurate disease identification, and making the solution more accessible to a broader audience.

Keywords: Aquaculture health, fish disease AI, automated diagnosis, image-based detection, machine learning fish, deep learning pathology, real-time monitoring, disease spread prevention, economic loss reduction, diagnostic accuracy, scalable solutions, rapid intervention.

1. INTRODUCTION

This research aims to develop an AI-driven system for detecting fish diseases in aquatic life. The system leverages machine learning, particularly deep learning algorithms such as Convolutional Neural Networks (CNN), to analyze images of fish and identify potential diseases based on visible symptoms. The model is trained on a large dataset of labeled fish images, ensuring high accuracy in disease classification. The web-based application allows users to upload fish images and receive predictions about possible diseases, providing real-time, automated diagnostics for the aquaculture industry. This AI-driven approach offers scalability, speed, and accuracy compared to traditional methods.

In traditional fish disease detection, methods like visual inspection and manual analysis are time-consuming, inaccurate, and heavily reliant on human expertise. These methods often lead to delays in diagnosing diseases, allowing them to spread and cause economic loss. Additionally, the process requires specialized personnel and facilities, making it difficult for small-scale aquaculture operations to effectively manage fish health. The lack of real-time and automated detection systems presents a significant challenge in maintaining the health and sustainability of aquatic life.

The motivation behind developing this research stems from the growing need for efficient, scalable, and accurate fish disease detection methods. With the increasing demand for seafood and the challenges of ensuring fish health, there is a critical need for rapid, automated solutions. The integration of AI into disease detection provides a means to address the inefficiencies of traditional methods, enabling real-time diagnoses, improving accuracy, and making the technology accessible to a wider range of aquaculture practitioners.

2. LITERATURE SURVEY

Yuhong Zheng, Zetian Fu, and Xiaoshuan Zhang [1] conducted a study on the design of a fish disease expert system, which was published in The Transaction of China Agricultural University in 2000. Their research focused on developing an expert system that could diagnose fish diseases using domain-specific knowledge. The system was designed to improve disease detection efficiency and provide accurate recommendations for fish health management. The expert system incorporated decision-making processes based on symptom analysis and knowledge-based inference.

Yuhong Zheng [2] further explored the development of a fish disease diagnosis expert system based on the web in 2000. This study, conducted at China Agricultural University, covered a detailed analysis of web-based expert systems for diagnosing fish diseases. The system aimed to provide remote access to fish disease diagnosis tools, allowing aquaculture farmers and researchers to detect diseases more efficiently. It included a user-friendly interface with a knowledge base that utilized expert consultations and past case studies.

Rustam et al. [3] performed a comparative analysis of supervised machine learning models for sentiment analysis on COVID-19 tweets. Their research evaluates models like Random Forest, Logistic Regression, and LSTM networks, focusing on their accuracy and efficiency in processing large datasets. They emphasize the need for balanced datasets to improve classification performance. Their approach to data balancing and model evaluation is relevant to VR user experience classification, where diverse interactions need to be analyzed in real-time.

I. Hatzilygeroudis, P. J. Vassilakos, and A. Tsakahdis [4] presented their research at the Medical Informatics Europe conference in 1997. Their study focused on medical informatics and expert system applications in the field of fish disease diagnosis. The system utilized artificial intelligence to improve diagnostic accuracy by analyzing symptoms and environmental conditions. The research emphasized the importance of integrating AI-driven decision support tools to enhance the reliability of disease detection systems.

Zetian Fu, Xiaoshuan Zhang, and Wen Jiwen [5] conducted a study on domain knowledge conceptualization in fish disease diagnosis in 2003. Their research, published in System Sciences and Comprehensive Studies in Agriculture, introduced an architecture for structuring domain knowledge in an expert system. The study highlighted the significance of knowledge representation techniques in improving expert system efficiency. Their conceptualization framework aimed to optimize the process of diagnosing fish diseases by organizing expert knowledge into a structured model.

Ding Wen, Wentong Li, Mary Li, and Yuhong Zheng [6] collaborated on the design of a fish disease expert system, which was published by ISIAIT in 2000. Their research focused on integrating multiple diagnostic methods into a unified expert system. The system utilized case-based reasoning and rule-based inference to enhance accuracy in disease detection. By combining different diagnostic approaches, the system aimed to provide a comprehensive solution for aquaculture disease management.

Daoliang Li, Zetian Fu, and Yanqing Duan [7] introduced ES-FDD, a web-based expert system for fish disease diagnosis, in 2002. Their study, published in Expert Systems with Applications, explored the use of web technologies to enhance accessibility and usability of fish disease diagnosis tools. ES-FDD was designed to provide online consultations and expert recommendations, enabling fish farmers to diagnose diseases in real time. The system incorporated AI algorithms to improve prediction accuracy and facilitate decision-making.

Daniel Zeldis [8] and Shawn Prescott analyzed problems and solutions in fish disease diagnosis programs in 2000. Their study, published in Aquacultural Engineering, addressed challenges such as data inconsistencies, environmental variations, and disease symptom overlap. They proposed solutions such as advanced imaging techniques and AI-driven models to improve diagnosis reliability. Their research emphasized the need for continuous improvements in fish disease detection systems.

3. PROPOSED METHODOLOGY

The proposed system aims to enhance the accuracy and efficiency of fish disease detection through the application of advanced machine learning techniques, specifically Convolutional Neural Networks (CNNs). The development process is structured as follows:
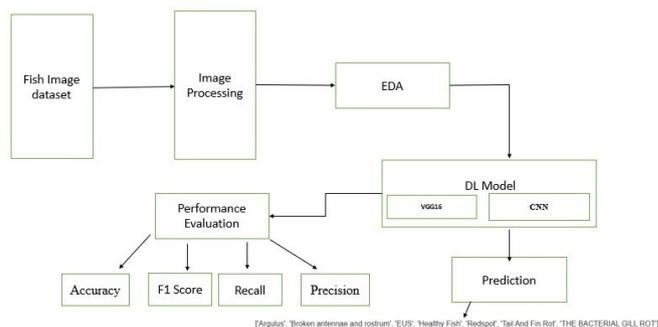

Figure 1: Block Diagram

The proposed system introduces a custom-designed CNN tailored specifically for fish disease detection. This model incorporates multiple convolutional layers to extract hierarchical features from input images, followed by pooling layers to reduce dimensionality and computational load. The architecture concludes with fully connected layers that perform classification based on the extracted features. The design aims to optimize accuracy and efficiency for the specific task of fish disease detection.

Applications:

- Real-time disease detection in aquaculture farms.

- Early diagnosis of fish diseases, preventing their spread.

- Automated classification of fish diseases based on image analysis.

- Cost-effective disease management for small-scale fish farms.

- Improved decision-making for fish health management.

- Integration with existing aquaculture systems for better monitoring.

- Enhanced fish disease research through data-driven insights.

- Training and educational tool for aquaculture professionals.

Advantages:

CNNs offer several advantages over traditional machine learning algorithms:

- **Automatic Feature Extraction**: CNNs autonomously learn relevant features from raw input data, eliminating the need for manual feature engineering.

- **Parameter Sharing**: Convolutional filters are applied across the entire input, reducing the number of parameters and computational load.

- **Translation Invariance**: CNNs can recognize objects regardless of their position in the image, enhancing robustness.

- **High Accuracy**: CNNs have achieved state-of-the-art performance in various image and video recognition tasks.

- **Robustness to Noise**: CNNs are robust to noise and distortion in the input data, which makes them highly effective in real-world applications.
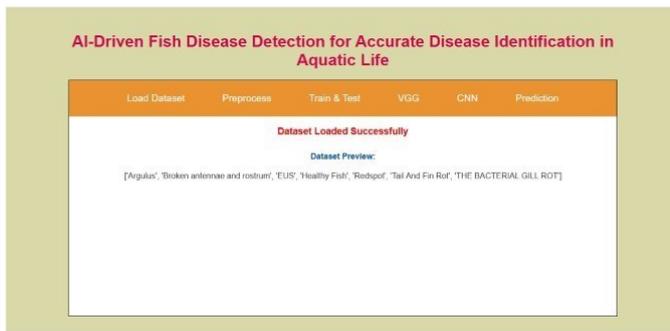
4. EXPERIMENTAL ANALYSIS

Figure 1: Home Page

This Figure 1 shows that user interface provides a user-friendly way to load, preview, and process image datasets for image classification using CNNs. The tab-based navigation guides the user through the typical machine learning workflow. The dataset preview allows users to verify the loaded data. The presence of "VGG" and "CNN" tabs suggests flexibility in model selection. The subsequent steps will likely involve data preprocessing, model training/evaluation, and finally, prediction on new images.
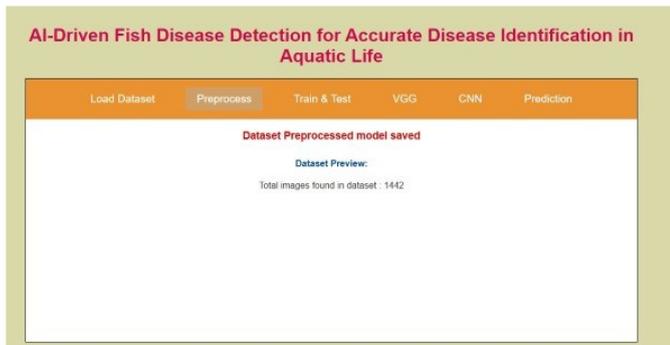

Figure2: Data Preprocessing

The above Figure 2 shows the successful completion of the image preprocessing stage. The application confirms that the dataset has been processed and a model related to this stage has been saved. The user can now proceed to the next steps of training and evaluating a CNN model for image classification.
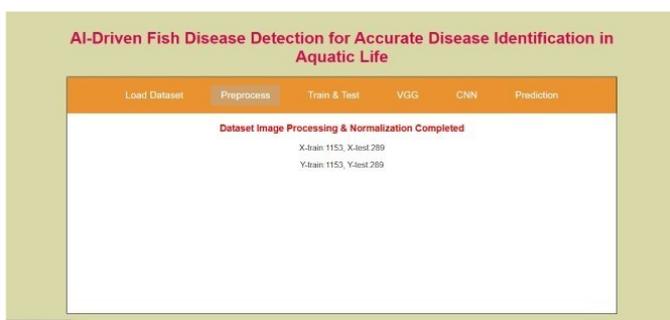

Figure 3: Train Test Splitting

"X-train 1153, X-test 289": This indicates that the dataset has been split into training and testing sets. "X-train" refers to the training images (1153), and "X-test" refers to the testing images (289). The "X" likely represents the image data itself.

"Y-train 1153, Y-test 289": This shows the corresponding labels (ground truth) for the training and testing sets. "Y-train" and "Y-test" represent the labels for the training and testing images, respectively.

Preprocessing Complete: The message clearly indicates that the preprocessing, including normalization, is done. Normalization is a crucial step to ensure that the pixel values are on a similar scale, which helps improve model training.

Train/Test Split: The data has been divided into training and testing sets. This split is essential for evaluating the model's performance on unseen data and preventing overfitting. The numbers suggest an 80/20 split (approximately), which is a common practice.

Data Ready for Training: With preprocessing and data splitting complete, the data is now ready for the next step, which is training and testing the model.

Figure 4: VGG16 Model Training

The Figure 4 shows the "Train & Test" stage of an image classification application using a VGG16 model.

Training Completed. Accuracy: 86.85%: This message indicates that the training process is finished, and the VGG16 model achieved an accuracy of 86.85% on the test set.

Prediction Plot: This section displays a confusion matrix, which is a visual representation of the model's performance.

VGG16 Algorithm Confusion matrix: The title clarifies that this is the confusion matrix for the VGG16 model.

Confusion Matrix Insights:

- The confusion matrix provides a detailed view of the model's performance for each class.

- The diagonal cells show that the model performs well at correctly classifying images for most classes, especially "GILL ROT" and "Healthy Fish."

- Misclassifications are visible in the off-diagonal cells. For example, there's some confusion between "Redspot" and "GILL ROT," and between "EUS" and "rostrum."
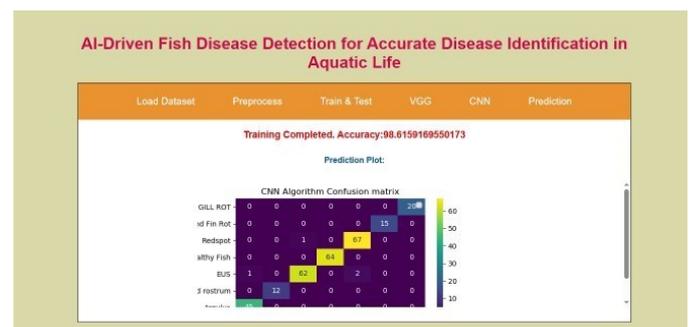

Figure 5: CNN Model Training

This Figure 5 shows the results of training and testing a Convolutional Neural Network (CNN) for image classification.

Training Completed. Accuracy: 98.61%: This message indicates the training is finished, and the CNN model achieved a high accuracy of 98.61% on the test set.

Prediction Plot: This section displays the confusion matrix.

CNN Algorithm Confusion matrix: This clarifies that the confusion matrix is for the CNN model.

Confusion Matrix Details:

- The model is particularly good at identifying "GILL ROT," "Fin Rot," "Redspot," and "Healthy Fish" with very high accuracy (as seen in the diagonal).

- There seems to be some confusion between "Redspot" and "GILL ROT," as well as between "EUS" and "rostrum."

- The last label (truncated) also shows some misclassifications.

Figure 6 : Predicted Output

The figure 6 shows the prediction results of an image classification model. Predictions: Healthy Fish": This text at the top indicates the model's prediction for the image. It's classifying the fish in the image as "Healthy Fish." Image: The main portion of the image shows a close-up view of a fish. And thus the image give the conclusion of the prediction of the diseases to reduce the misconception of the diseases.

5. CONCLUSION

The integration of machine learning, particularly Convolutional Neural Networks (CNNs), into fish disease detection has significantly advanced the field of aquaculture. By leveraging image-based analysis, these technologies enable early and accurate identification of various fish diseases, thereby enhancing fish health management and reducing economic losses. Studies have demonstrated that deep learning models, such as CNNs, outperform traditional machine learning methods in classifying fish diseases, offering higher accuracy and efficiency. The application of CNNs in fish disease detection involves training models on extensive datasets comprising images of healthy and diseased fish. These models learn to recognize patterns and anomalies in fish images, facilitating the identification of diseases like Argulus, Redspot, and Tail and Fin Rot. The ability to process and analyze large volumes of image data allows for rapid and precise disease detection, which is crucial for timely intervention and treatment.

Furthermore, the use of CNNs in this domain contributes to the development of automated systems that can operate continuously, providing consistent monitoring of fish health. This automation reduces the reliance on manual inspections, which can be time-consuming and prone to human error. Additionally, the integration of CNNs with other technologies, such as Internet of Things (IoT) devices and sensors, can lead to the creation of comprehensive monitoring systems that offer real-time data on fish health and environmental conditions.

One promising direction is the integration of machine learning models with Internet of Things (IoT) devices. By embedding sensors in aquaculture environments, real-time data on water quality, temperature, and other critical parameters can be collected and analyzed. This integration would enable proactive disease management, allowing for immediate responses to environmental changes that may affect fish health.

Another area of potential growth is the enhancement of datasets used for training machine learning models. Currently, many datasets are limited in size and diversity, which can affect the generalization capabilities of models. Expanding these datasets to include a broader range of fish species, environmental conditions, and disease variations will improve the robustness and accuracy of detection systems. Advancements in computational power and algorithm optimization also present opportunities for more efficient and faster disease detection. Implementing transfer learning techniques, where models trained on large, diverse datasets are fine-tuned for specific applications, can reduce training times and improve performance. Furthermore, the development of user-friendly interfaces and mobile applications can facilitate the adoption of these technologies by fish farmers. Providing accessible tools for disease detection and management will empower farmers to make informed decisions, leading to healthier fish populations and more sustainable aquaculture practices.

REFERENCES

[1] Yuhong Zheng, Zetian Fu and Xiaoshuan Zhang, "Design of fish disease expert system", The transaction of China Agricultural University, vol. 5, no. 6, pp. 94-97, 2000.

[2] Yuhong Zheng, 2000b. Study on fish disease diagnosis expert system based on Web, Beijing:China Agricultural University, pp. 1-30, 2000.

[3] Chengzhi Wang, Shaotao Huang and Rongxing Ji, Fish disease diagnosis system – 'Fish doctor', The transaction of Jimei University, vol. 3, 1997.

[4] I. Hatzilygeroudis, P. J. Vassilakos and A. Tsakahdis, *1997. Published in the Proceedings of the Medical Informatics Europe'97 (MIE'97)*, pp. 259-299, May 1997.

[5] Zetian Fu, Xiaoshuan Zhang and Wen jiwen, "Structure on domain knowledge conceptualization architecture in a fish disease diagnosis expert system", *System Sciences and Comprehensive Studies in Agriculture*, vol. 19, no. 3, pp. 23-25, 2003.

[6] Ding Wen, Wentong Li, Mary Li and Yuhong Zheng, Diagnosis of fish disease and design of fish disease expert system, Beijing:ISIAIT, pp. 183-187, 2000.

[7] Daoliang Li, Zetian Fu and Yanqing Duan, "2002. ES-FDD: a web-based expert system for fish disease diagnosis", *Expert Systems with Applications*, vol. 23, pp. 311-320, 2002.

[8]   Daniel Zeldis and Shawn Prescott, "2000. Fish disease diagnosis program-problems and some solutions", *Aquacultural Engineering*, vol. 23, pp. 3-11, 2000.

[9]   Noraini Hasan, Shafaf Ibrahim, Anis Aqilah Azlan , "Fish Diseases Detection Using Convolutional Neural Network (CNN)." Int. J. Nonlinear Anal. Appl, vol. 13, 2022.

[10]   Md. Jueal Mia, Rafat Bin Mahmud , Md. Safein Sadad , Hafiz Al Asad , Rafat Hossain "An in-depth automated approach for fish disease recognition". Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 9, Oct. 2022.

[11]   Md Shoaib Ahmed , Tanjim Taharat Aurpa, Md. Abul Kalam Azad ."Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture". Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 8, Sept. 2022