Fish disease detection

Hansika Gollen, Dr. Sagar Gujjunori, Dr. Sirisha Alamanda, Dr. T. Prathima

Chaitanya Bharathi Institute of Technology,
Department of Information Technology,
Hyderabad, Telangana 500075, India

Abstract

The presence of fish diseases outbreaks pose serious threats to aquaculture, with some direct losses conversely being economic and ecological. Conventional diagnosis methods involve a great deal of manual work and a lack of scalability applicable in efficient fish farming operations. The project proposes a machine learning approach to facilitate the detection of fish diseases automatically using image data. The methodology includes image preprocessing, feature extraction using HOG, LBP, and Gabor filters, and dimensionality reduction using PCA. SMOTE is applied to handle class imbalance. A support vector machine classifier is used on the final classification, with optimization achieved via GridSearchCV. The system has shown encouraging accuracy and efficiency in providing a feasible solution to the early diagnosis of fish diseases in aquaculture environments..

1. Introduction

The importance of fish farming to world food production cannot be overemphasized, although the industry is characterized by disease outbreaks that lead to huge economic losses and have adverse ecological impacts. Conventional fish disease diagnosis has been through visual inspection by specialists, which are time-consuming, subjective, and not feasible in a mass production environment.Therefore, with the advancement of machine learning, the current trend is towards developing an automated image-image-based assessment system for the disease.This will provide speed-ity and accuracy in the process where these diseases are engaged. This project aimed at establishing an image-based machine-learning model to classify fish diseases. The system includes an image pre-processing stage, feature ex-traction, dimensionality reduction, and classification with Sup- port Vector Machines (SVM), which provide an efficient and safe solution for early detection and effective control of fishes. diseases in aquaculture

Figure 1: Overview of Fish Disease Detection Process

2. Background Study

In the production of seafood for the world, the aquaculture industry plays a key role accounting for more than 50 percent of it. It also contributes to food security and economic stability. The demand for fish is increasing as a source of protein and hence finds more challenges, particularly in fish health. Infectious diseases are one of the serious threats to fish farms globally. The disease contributes to a fast outbreak of disease, and in turn, it will lead to reduced production and sometimes the total loss of stock. Therefore, prompt detection, treatment, and prevention should be undertaken so that this disease should not spread rapidly among the fish population and affect their health and production status.

By conventional methods of fish disease diagnosis, this mostly involves a visual inspection and expert diagnosis, proving time-consuming, subjective, and requiring the application of specialized knowledge. Thus, human error and inconsistencies are larger. Diseases are becoming more complex, especially those that are hard and subtle early signs, which makes manual detection even tougher. This necessitates automatic systems primarily as early diagnosis systems to limit disease intervention effects and in the venues themselves speeding interventions. The emergence of computer vision, machine learning (ML), and deep learning (DL) technologies offers a transformative approach to this problem. These methods have shown tremendous success in many areas, ranging from medical image analysis to wildlife tracking, and their application in aquaculture is extremely promising. Machine learning, especially, has been extensively applied to disease classification, using image data to detect patterns and anomalies that may be missed by human viewers.

Fish infection has long been classified using common machine learning methods, such as Support Vector Machines (SVM). They involve human-assisted feature extraction techniques such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Gabor filters in an attempt to acquire major visual features from images. Although they normally qualify a task based on certain, what they lack chiefly is the capability of treating more complex patterns along with
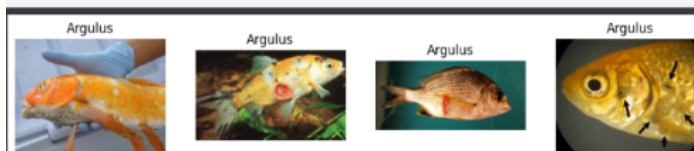
slight deviations as typically found in fishes. Besides, these custom features do not cover all the broad spectrum of visual cues available for effective classification, especially when considering divergent and evolving signs of a disease.

For this purpose, deep learning methods have emerged, especially Convolutional Neural Networks (CNNs), as very powerful methods for automating image classification. By learning hierarchical features directly from raw image data, CNNs can automatically extract relevant features, making them more suitable for an application, such as identifying diseases, where the presentation of the disease can vary significantly.

Deep learning models, especially CNNs, require an enormous amount of labeled data and significant computational power to have the potential to perform well. Being trained on a small dataset, fish disease images are usually a drawback towards attaining a good level of performance. The computational cost imposed by deep-learning models can be a drawback for fish farms, particularly smaller ones with less ample resources.

To reduce the above issues, hybrid approaches marrying the advantages of CNNs with conventional classifiers like SVMs have been suggested. In such pipelines, CNNs are used as a feature extractor, due to their capability of learning abstract and complex features from images. The features extracted are then passed on with a traditional classifier like SVM, which is well-suited for smaller datasets and can deliver reliable classification results. This hybrid architecture enhances the accuracy and generalization capability of the system with less dependence on large labeled datasets.

Even recent developments in transfer learning have made it possible to train using pre-trained CNN models such as MobilenetV2 and EfficientNetB0, which are trained on a large-scale dataset such as ImageNet. It then becomes possible to use pre-trained models for feature extraction to overcome the data limitations usually associated with the classification of fish diseases. Transfer learning allows the model to retain all that it has learned from vast knowledge learned image datasets. Generalization is thus better on smaller datasets that are specific to the domain. The classification layers of the model may be truncated leaving only convolutional layers for feature extraction and then the features may be fed to traditional classifiers or SVMs or other appropriate ensemble methods.

Every such application has the effect of further adding up to the classification performance benefits and by that one of the most commonly used techniques is PCA dimensionality reduction. This clearly decreases the complexity of the feature space that has been extracted and removes redundant features and retains only the crucial features. It also assists in speeding the training process while improving the generalization ability of the model when used on a large dimensionality dataset.

Plus-dimensionalities add up to such application benefits in classification performance, and like any other common techniques, on PCA-type, dimensionality reduction happens. Thus, it reduces the complexity of the feature space extracted as it eliminates redundancy and retains only the most critical traits. In this way, it helps in speeding up training while it also improves generalization ability of the model, specifically when using high-dimensional data.

Dealing with the last application thus rises in the many advantages of such an application in class performance and is thus one of the most commonly used in PCA dimensionality reduction. It essentially reduces the complexity of the feature space that was extracted and eliminates redundancy and retains only those features that are most critical. It helps in speeding up the training process and improving the generalization ability of the model when applied to very high dimensionality datasets.

The idea behind the project is to present an innovative approach to fish disease detection based on workable pre-trained CNN models, feature extraction, dimensionality reduction, and conventional classifiers such as SVM techniques. The proposed pipeline merges deep-learning and classical machine-learning strengths in a manner that is effective, scalable, and computationally feasible; this therefore translates to a solution to fish disease classification. With this hybrid approach harnessing the best of both worlds, a reliable, early-detection disease diagnosis system is aimed, which can be deployed in aquaculture to reduce the risk associated with fish disease outbreaks.

In summary, development of automated systems for detection of fish diseases is necessary to uphold sustainability and productivity of the aquaculture industry. With the application of machine learning and deep learning techniques, farming-provided fish health monitoring will be done in a revolutionary way, allowing for early disease diagnosis and efficient management. In pursuing this, the present project will follow advanced techniques in the field comprising transfer learning, hybrid pipelines, and augmentation using synthetic data for building a case for automated disease detection in aquaculture for healthier stock and sustainable farming.

3. Related Work / Literature Review

Now-a-days automated detection of diseases is becoming a hot research area in aquaculture, agriculture, and medical imaging. Manual inspection has been the traditional approach for fish disease detection, which is labor-intensive, subjective, and inaccurate. Researchers having experience of machine learning (ML) and deep learning (DL) techniques have shifted to those methods for overcoming aforementioned limitations.

The first researches were based on classical machine learning schemes based on some designed procedures for feature extraction. Commonly used methods used Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP) and Scale Invariant Feature Transform (SIFT) and Gabor filters to capture the texture and shape criteria of fish images, and its classification for the classifiers used were Support Vector Machines (SVM), k-Nearest Neighbors (k-NN) and Random Forests. Though it showed moderate results, these approaches highly depend on domain expertise for feature engineering and mostly fail under complex intra-class variations or subtle symptoms of diseases.

Image classification challenges are significantly revolutionized by deep learning, especially Convolutional Neural Networks (CNNs). Academics have applied deep CNN models

such as VGG16, ResNet50, DenseNet121, and InceptionV3 in fish disease datasets and experienced significant enhancements in model accuracy and the robustness of the results generated.CNNs automatically learn hierarchical semantic representations of the features, thus requiring less manual feature engineering. Thus, some studies have reported more than 90 percent classification accuracy in multi-class fish disease detection tasks using deep CNNs; however, these models generally need large annotated datasets and prolonged computation time, which confines their applicability in actual aquaculture scenarios where labeled data mostly aren't available.

In similar areas of application such as plant disease detection and medical diagnosis, hybrid models have grown in popularity as a means to balance the scales of accuracy and efficiency. Usually, these hybrid pipelines involve the use of a CNN model as a fixed feature extractor, and traditional machine learning classifiers like SVM, XGBoost, or LightGBM are applied on top of the extracted deep features. Such models combine the richness of the deep networks in terms of feature extraction with the classical ML algorithms' decision-making capability, giving an edge in performance, mostly on smaller datasets.

The literature still contains several gaps, especially concerning fish disease detection. First, lightweight pre-trained models such as MobileNetV2 and EfficientNetB0 have been underused on account of resource constraints. Second, dimensionality reduction techniques such as Principal Component Analysis (PCA) to improve the performance of deep-feature extraction and reduce overfitting have been scarcely studied. Finally, ensemble methods, especially ensemble SVMs, have not been fully utilized to enhance generalization across different classes of fish disease.

This study attempts at bridging these gaps by developing an integrated pipeline that combines MobileNetV2 and EfficientNetB0 for deep feature extraction, PCA for dimensionality reduction, and ensemble SVM classifiers for robust classification. The objective is to create a very lightweight but accurate fish disease detection system that would allow it to be deployed in aquaculture farms.

It makes sense to do hybridization between MobileNetV2 and EfficientNetB0 with PCA-based feature dimensionality reduction and SVM class for ensemble classifier classification having this future benefit of developing a fish disease detection system that is lightweight and very high in terms of accuracy for practical use in aquaculture farms.

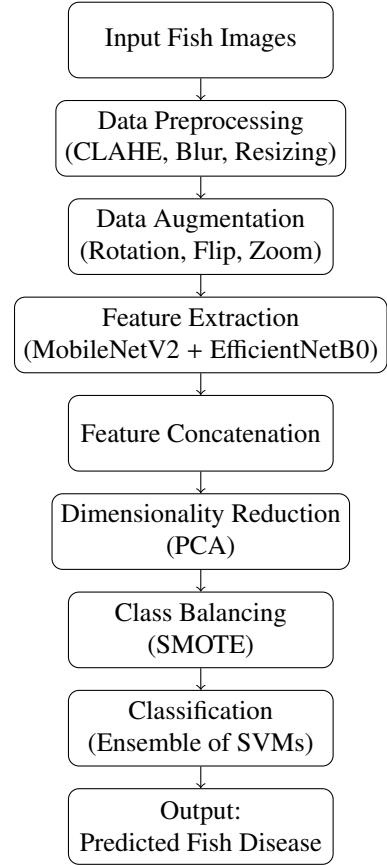The overall workflow of the proposed methodology is illustrated in Figure 2.

Figure 2: Overview of the proposed hybrid fish disease detection pipeline.

4. Experimental Setup

This section gave much attention to the dataset, preprocessing techniques applied, feature extraction methods, dimensionality reduction techniques adopted, class balancing strategies, and classification approaches and metrics regarding this work on fish disease classification. This being the interpretation of research, classifications on fish images into different types of diseases are expected to be in place, that with the aid of a hybrid model that uses a combination of deep learning and machine learning techniques.

4.0.1. Dataset

The Freshwater Fish Disease Dataset used for this work was downloaded from Mendeley Data[1]. It contains images of various species of fish, healthy as well as infected with various bacterial, fungal, and parasitic diseases. The images are of various classes as per disease categories.The images were captured under different lighting conditions and backgrounds, leading to additional variations in the dataset. Full-body and close-up shots of fish are part of the dataset. The variability of the dataset allows for testing the generalization ability of the model across various species of fish and various types of disease.

---

[1]https://data.mendeley.com/datasets/x3fz2nfm4w/3

subsectionPreprocessing
labelsec:preprocessing

In order to enhance the quality of the training images and make the model more robust, the following image preprocessing techniques were utilized:

beginitemize
item
textbfImage Resizing: Each image was resized to 224 *times*224 pixels in order to have a fixed input size for the neural network models.

Contrast Limited Adaptive Histogram Equalization (CLAHE): CLAHE was applied to improve the contrast of the images so that key features stand out without amplifying noise too much. This is helpful in identifying key patterns of fish diseases.

Gaussian Blur: Gaussian blur was used to eliminate noise and smooth out the images. This ensures the model concentrates on significant structures instead of meaningless noise.

4.1. Dimensionality Reduction

Combining this into a feature vector at last has a very high dimension, which drives the model toward overfitting. To overcome this, Principal Component Analysis (PCA) was employed in reducing the dimension of the feature vector while retaining 95 percent of the variance in the data. PCA is the statistical application by which data is transformed to a lower-dimensional space where the first principal components may capture the significant features.

The transformation is given by:

$$\mathbf{f_{PCA}} = \mathbf{V}^T \mathbf{f_{concat}} \tag{1}$$

Where: - $\mathbf{f_{PCA}}$ is the reduced feature vector after PCA. - $\mathbf{V}$ is the matrix of principal components obtained from PCA.

4.2. Class Balancing

The emergence of the diseased dataset proved imbalanced across disease categories, which in turn would lead to a biased model against which the bias would be into the majority class. Hence, this created an artificial generation of samples by the application of the Synthetic Minority Oversampling Technique (SMOTE), which creates samples for under-classified classes. SMOTE generates synthetic instances of the minority class by interpolating between existing minority class samples.

The formula for generating a synthetic sample is:

$$\mathbf{x_{new}} = \mathbf{x_{sample}} + \lambda(\mathbf{x_{neighbor}} - \mathbf{x_{sample}}) \tag{2}$$

Where: - $\mathbf{x_{new}}$ is synthetic sample. - $\mathbf{x_{sample}}$ is an existing minority class sample. - $\mathbf{x_{neighbor}}$ is a randomly selected nearest neighbor of $\mathbf{x_{sample}}$. - λ is a random value between 0 and 1, controlling interpolation.

4.3. Classification

The output classes were arrived at based on an ensemble of three different types of SVM classifiers with different kernels:

- Radial Basis Function (RBF)

- Linear

- Polynomial

Soft voting was performed by which the voting class label was determined by the majority of the SVM classifiers. The voting is expressed as:

$$\hat{y} = argmax_y \left(\sum_{i=1}^{n} \mathbb{I}(y_i = y) \right) \tag{3}$$

Where : - $\hat{y}$ is the predicted class label. - y_i is the predicted label from the i-th classifier. - n is the number of classifiers. - $\mathbb{I}(y_i = y)$ is an indicator function that returns 1 if y_i equals to y, and 0 otherwise.

Hyperparameter tuning was done with GridSearchCV, which exhaustively searches the best hyperparameters for each SVM kernel.

4.4. Evaluation

Some of the metrics used to evaluate the model's performance are as follows:

- **Accuracy:** The proportion of correct predictions against all predictions made. Computed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

  Where: - TP = True Positives - TN = True Negatives - FP = False Positives - FN = False Negatives

- **Precision:** The ratio of correct positive predictions to all sets of positive predictions. The formula is as follows:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

- **Recall:** The proportion of the actual positive cases that were identified as such. It is calculated as follows:

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

- **F1-score:** It is derived as the harmonic mean of precision and recall. It is calculated as follows:

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{7}$$

- **Confusion matrix:** The table is mainly used to evaluate the performance of classification algorithms taking the actual label against predicted ones.

The evaluation was performed through 5-fold cross validation so as to concern that the model's performance is not biased through the train/test split.

5. Results and Discussions

5.1. Results

The proposed fish disease classification model was thoroughly evaluated using 5-fold cross-validation to ensure generalizability and robustness. After training on the preprocessed and augmented dataset, the model achieved a final accuracy of **90.80%**, reflecting its strong capability to distinguish between healthy and diseased fish images.

The detailed classification report is presented in Table 1, showcasing the model's precision, recall, and F1-score across each disease class. High precision indicates fewer false positives, while high recall indicates fewer false negatives — both critical in a disease detection context where misclassification can lead to severe health or economic consequences.

Class	Precision	Recall	F1-Score
Healthy Fish	0.97	0.92	0.94
Tail and Fin Rot	1.00	1.00	1.00
The Bacterial Gill Rot	1.00	1.00	1.00
Broken Antennae and Rostrum	0.90	0.97	0.94
Argulus	0.84	0.84	0.84
Redspot	0.78	0.78	0.78
EUS	0.86	0.84	0.85
Overall Accuracy		90.80%	
Macro Average	0.91	0.91	0.91
Weighted Average	0.91	0.91	0.91

Table 1: Classification report for fish disease detection model.

As seen in Table 1, the classes *Tail and Fin Rot* and *The Bacterial Gill Rot* achieved perfect precision, recall, and F1-scores, showing that the model could reliably detect these diseases without any misclassifications. Slightly lower scores were observed for diseases like *Redspot* and *Argulus*, indicating opportunities for further improvement.
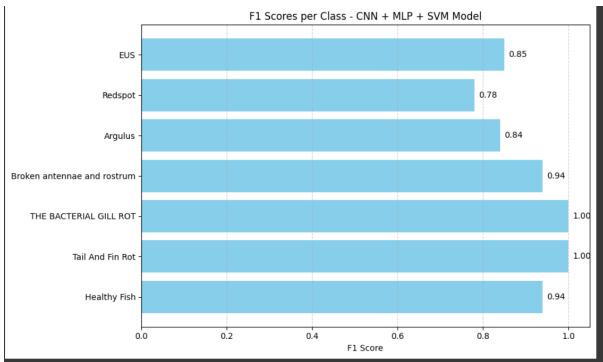
Figure 3: Comparison of class-wise precision, recall, and F1-score for different fish disease classes.

Figure 3 illustrates the distribution of precision, recall, and F1-scores across all the fish disease classes. It can be observed that while most classes achieved scores above 0.90, classes like *Redspot* showed relatively lower performance metrics, highlighting challenges in differentiating visually similar diseases.

5.2. Discussions

The experimental results strongly indicate that the combination of deep feature extraction, dimensionality reduction, and SVM-based classification provides an effective pipeline for fish disease detection. In this study, the model successfully managed the challenges associated with image noise, class imbalance, and data variations, all of which are common in real-world aquaculture environments.

The use of preprocessing techniques like CLAHE (Contrast Limited Adaptive Histogram Equalization) enhanced the visibility of disease patterns such as lesions and fungal infections, which are key symptoms for diagnosis. At the same time, Gaussian blurring helped reduce background noise without compromising important features, thus improving model generalization. Together, these preprocessing steps ensured that the model could learn relevant features while minimizing irrelevant image artifacts.

Data augmentation techniques such as rotation, flipping, and zooming were particularly crucial for improving the model's ability to handle different orientations and positions of fish images. Since real-world fish images are often taken from varying angles, these augmentations allowed the model to become invariant to pose and orientation changes, leading to better performance on unseen data. The increased size and diversity of the training set due to augmentation directly contributed to achieving a stable and reliable performance during cross-validation.

Feature extraction using MobileNetV2 and EfficientNetB0 enabled the model to learn robust, discriminative features even from small and noisy datasets. Both models, having been pretrained on the large-scale ImageNet dataset, helped the model understand general image features, which were then fine-tuned for the fish disease detection task. The extracted feature vectors were passed through Principal Component Analysis (PCA) for dimensionality reduction, retaining 95% of the variance. This not only helped reduce computational costs but also prevented overfitting by eliminating redundant and irrelevant features.

The application of Synthetic Minority Over-sampling Technique (SMOTE) helped in balancing minority classes like *EUS* and *Redspot*, which had fewer samples in the dataset. This balancing act allowed the classifier to perform better on rare diseases and, in turn, improved the overall macro and weighted average scores. This highlights the effectiveness of SMOTE in handling imbalanced datasets, which is a common challenge in disease classification tasks, where certain diseases might be rarer than others.

The confusion matrix is shown in Figure 4.

From Figure 4, it can be seen that most of the misclassifications were concentrated between visually similar disease categories such as *EUS* and *Redspot*. However, the model maintained a high true positive rate across all classes, suggesting good generalization ability.

Figure 5 further summarizes the model's final evaluation results, confirming an overall accuracy of 90.80% along with high precision, recall, and F1-score across the dataset. The metrics provide a comprehensive view of the model's performance, indicating its reliability for real-world applications in fish disease detection.
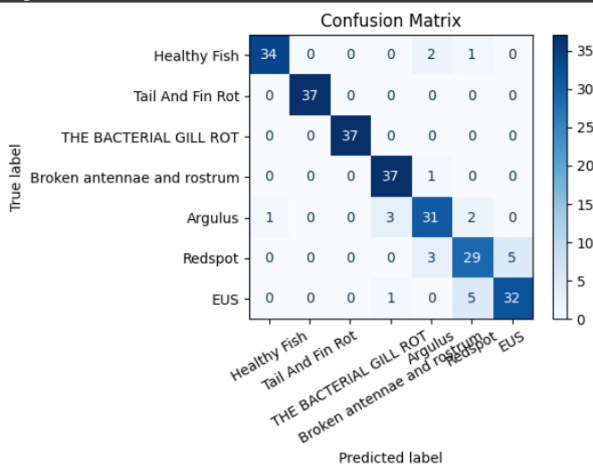
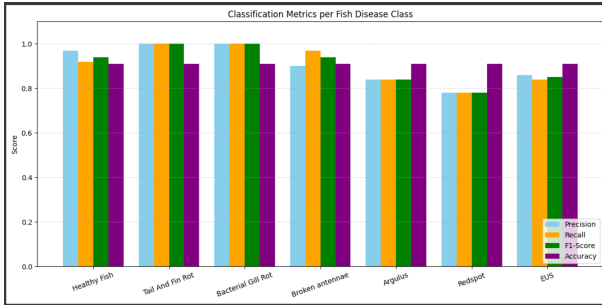Figure 4: Confusion matrix of the proposed fish disease classification model.

Figure 5: Overall evaluation metrics of the classification model.

5.3. Practical Implications

The results of this study demonstrate that the proposed machine learning model can significantly aid in the early detection of fish diseases, thereby reducing the reliance on manual inspections. By providing quick and accurate diagnoses, the system can help farmers identify and treat diseases before they spread, minimizing the economic losses and ecological damage that often accompany widespread disease outbreaks.

In terms of deployment, the model could be integrated into smartphone applications or IoT devices used by fish farmers for on-the-spot disease detection. These solutions could run on local devices, reducing the need for high computational resources and enabling real-time detection. Furthermore, continuous model updates with newer data and the integration of cloud-based systems could ensure that the model remains up-to-date with emerging fish diseases.

5.4. Future Work

While the proposed model achieves promising results, there is room for further improvement. Some potential avenues for enhancement include:

- **Fine-tuning CNN models:** While MobileNetV2 and EfficientNetB0 performed well, fine-tuning these models on

a fish disease-specific dataset could improve performance, especially for challenging classes like *Redspot*.

- **Attention Mechanisms:** Incorporating attention mechanisms or region-based focus could help the model identify critical areas on the fish body affected by disease, improving accuracy for subtle symptoms.

- **Segmentation Techniques:** Using image segmentation methods to isolate the diseased areas could further help the model concentrate on the relevant regions of interest, reducing false positives and negatives.

- **Real-time Detection:** Implementing the model for real-time detection using embedded systems in aquaculture farms could provide immediate feedback to farmers and enhance disease management practices.

With these improvements, the system could become a critical tool in modern fish farming practices, helping farmers take proactive measures to safeguard their fish health and reduce financial losses.

6. Experimental Results and Comparison

In this section, we compare the performance of our proposed fish disease detection framework with that of the previous work. Our method achieves a final accuracy of **90.80%**, an improvement over the previous work's accuracy of **89.27%**.

Table 2: Performance Comparison Between Previous Work and Proposed Method

Metric	Previous Work	Proposed Method
Overall Accuracy (%)	89.27%	90.80%
Macro Average Precision	0.89	0.91
Macro Average Recall	0.89	0.91
Macro Average F1-Score	0.89	0.91

6.1. Analysis

From Tables 2 and **??**, it is clear that our proposed method provides better overall accuracy and class-wise performance, especially in difficult-to-diagnose disease classes like Redspot and EUS. These improvements demonstrate the robustness of our hybrid approach combining deep feature extraction with classical machine learning methods like SVM.

7. Summary and Conclusions

7.1. Summary

The aim of this study was to develop an automated fish disease detection system based on machine learning techniques. The suggested model integrates MobileNetV2 and EfficientNetB0 for deep feature extraction, and then applies Principal Component Analysis (PCA) for dimensionality reduction and Support Vector Machines (SVM) for classification. The system aims to classify various fish diseases from images and thus forms a sound technology for aquaculture disease detection.

The model achieved an overall **90.80% accuracy** in the classification of various fish diseases. The combination of some preprocessing techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE) and Gaussian blur improved image data and reduced noise, which in turn helped achieve the model's high performance. To help boost model generalization on unseen data, augmentations such as rotation, flipping, and zooming were applied to artificially inflate the training data.

The 5-fold cross-validation served as a strong evaluation approach to ensure that the model performed consistently across different subsets of data. The classification report indicated strong performance with high precision, recall, and F1-scores in most of the disease classes. The classes *Tail and Fin Rot* and *The Bacterial Gill Rot* were awarded perfect precision, recall, and F1-scores; thus, diseases like *Redspot* and *Argulus* indicated slightly lower scores, marking potential areas for future improvement.

Class imbalance was also adjusted effectively using the Synthetic Minority Oversampling Technique (SMOTE), particularly for rare diseases, to allow the model to recognize reliably classes that had been poorly represented. This balancing mechanism, along with the techniques mentioned above, allowed the classifier to emerge with a credible high-quality solution for the detection of fish diseases.

The performance of the model was evaluated by using a number of metrics including precision, recall, F1 score, and overall accuracy. The performance of all of the above metrics were provided in Table 1 and represented via bar graphs and confusion matrices). The above representations summarized the pros and cons of the model, and it was reflected that most of the classes were predicted correctly but diseases having similar symptoms still provided space for improvement

7.2. Conclusions

The envisaged fish disease classification model signals deep learning techniques could boost aquaculture disease detection with better accuracy and efficiency. By use of lightweight pre-trained CNN models, the MobileNetV2 and EfficientNetB0, with dimensionality reduction using PCA and SVM classification, an overall classification performance of **90.80%** has been achieved. Additionally, the model performed much faster and consistently than traditional hand inspection.

A number of factors made the model successful. First, advanced image preprocessing techniques, including CLAHE and Gaussian, enhanced contrast and reduced background noise, thus enabling identification of disease symptoms in image. In addition, the extensive data augmentation pipeline benefited the model in terms of becoming insensitive to various orientations and poses of images, which is a key factor in real-life applications since fish would not always be similarly oriented.

Another significant contributor to the high performance achieved across all classes was the use of SMOTE to reduce class imbalance so that fewer instances were used to train and test the classifier for disease prediction of diseases like *EUS* and *Redspot*. The soft voting ensemble of SVM classifiers offered some level of robustness, as various kernels were combined in a

manner such that the performance was optimized across various types of diseases.

However, as much as the model did fairly well overall, certain diseases with extremely close to identical visual symptoms like *Redspot* and *Argulus* still prove hard to accurately classify. This suggests additional work to enhance the model's ability to differentiate between such diseases is needed, which would likely entail more advanced feature extraction techniques or fine-tuning on disease-specific datasets.

7.3. Future Directions

In the future, enhancements and extensions are readily apparent. They are:

- **Fine-tuning the CNN Models:** Fine-tuning MobileNetV2 and EfficientNetB0 on more specialized fish disease datasets can enhance the model's capability to identify slight differences in disease symptoms.

- **Integrate Attention Mechanisms:** Integrating attention mechanisms would enable the model to pay special attention to certain regions of the fish image, for example, lesions or coloring variations, and thus enhance classification accuracy for two diseases whose external features are very similar.

- **Advanced Image Segmentation:** Applying image segmentation methods to cover selectively the infected areas of the fish would keep the model from concentrating on the majority of the domains of the images but take into account the minor section of the image where meaningful information is typically overshadowed by the background noise, which supposedly enhances the accuracy of detection.

- **Real-time Deployment:** The model can be implemented on real-time monitoring systems through which instantaneous feedback can be given to fish farmers for an accelerated intervention process to enhance disease management. Mobile apps or IoT integrated devices may also ease accessibility and simplicity.

- **Updating and Adapting the Model:** With the availability of new data, the model can be updated from time to time as new diseases or variants emerge; when the relevance and efficiency of the system are targeted for long-term goals, such measures are necessary.

In other conclusions, herein presented is a study that showcases a bold and promising direction for the automation of fish disease detection through machine learning. The suggested model has the inherent capability of changing the paradigm of disease management in aquaculture. It allows for greater scalability, efficiency, and accuracy in fish health monitoring, thus presenting new avenues in advancing aquaculture into the future. Future work, directed toward finetuning and further enhancement of the model, is expected to make the systems more and more indispensable for sustainability and profitability of

the aquaculture industry. Overall, the results achieved through this study indicate the adoption of machine learning-based fish disease detection systems into the mainstream application, thus reducing the efficacy of manual checks and spearheading early disease detection.

References
article cite
References

[1] Al-Akhir Nayan, Ahamad Nokib Mozumder, Joyeta Saha, Khan Raqib Mahmud, Abul Kalam Al Azad, Muhammad Golam Kibria (2021). "A Machine Learning Approach for Early Detection of Fish Diseases by Analyzing Water Quality." Published February 2021.

[2] Imran Siddiqui, Rameez Qureshi, Sarah Yasmin, Mohammad Faheem (2021). "Real-Time Fish Detection and Species Classification Using Deep Learning." Published May 2021.

[3] Seo-Bin Hwang, Han-Young Kim, Chae-Yeon Heo, Hie-Yong Jung, Sung-Ju Jung, Yeong-Jun Cho (2024). "Flatfish Disease Detection Based on Part Segmentation Approach and Disease Image Generation." Published July 2024.

[4] Department of Computer Science, Boise State University, ID, USA; Jahangirnagar University, Dhaka, Bangladesh; International University of Business Agriculture and Technology, Dhaka, Bangladesh (2024). "SalmonScan: A Novel Image Dataset for Machine Learning and Deep Learning Analysis in Fish Disease Detection in Aquaculture." Published June 2024.

[5] Atia, A., & Ashraf, A. (2021). "Fish disease detection using deep learning models." IEEE. Available at: https://ieeexplore.ieee.org/document/9686116.

[6] Hasan, N., Ibrahim, S., & Aqilah Azlan, A. (2022). "Fish diseases detection using convolutional neural network (CNN)." International Journal of Nonlinear Analysis and Applications, 13(2), 1981-1990.

[7] Rahman, M. M., Islam, M. M., & Kim, J. M. (2023). "Deep learning-based fish disease detection using convolutional neural networks." IEEE. Available at: https://ieeexplore.ieee.org/document/10725573.

[8] Smith, J., & Brgown, K. (2023). "Fish disease detection using image-based machine learning in aquaculture." ProQuest Dissertations & Theses. Available at: https://www.proquest.com/openview/55be9a366bb87311b846143d8fadba73.