

RESEARCH ARTICLE



An Intelligent Scheme of Automatic Fish Disease Recognition Using Optimized Deep Features

Md. Meharaj Uddin¹ , Ummatun Nahar Oishi¹ , Arslan Munir^{2,*}, Muhammad Minoar Hossain²,
Md. Khabir Uddin Ahamed³ and Md. Sadiq Iqbal¹

¹Department of Computer Science and Engineering, Bangladesh University, Bangladesh

²Department of Electrical Engineering and Computer Science, Florida Atlantic University, USA

³Department of Computer Science and Engineering, Bangamata Sheikh Fojilatunnesa Mujib Science and Technology University, Bangladesh

Abstract: Aquaculture involves the breeding of aquatic organisms for maintenance and conservation purposes. Detecting fish diseases remains a major challenge in aquaculture because of the diversity of fish species and the influence of environmental factors. This study has implications for improving fish disease detection in aquaculture. In this study, a dataset consisting of healthy and diseased fish images was collected from various sources. Several convolutional neural network models, including VGG16, VGG19, ResNet50, Inception V3, and Xception, were employed for automated feature extraction. Then the extracted features from each model were assessed using different machine learning (ML) classifiers, both with and without feature optimization for fish disease identification. For feature optimization, principal component analysis and linear discriminant analysis (LDA) were investigated. To further improve performance, a soft voting ensemble classifier was utilized by combining the outputs of multiple ML classifiers, thereby taking advantage of their individual strengths. The combination of Xception, LDA, and the voting ensemble achieved the highest accuracy of 98.3% for fish disease detection, and this outcome outperforms existing methods.

Keywords: fish disease, food supply, image processing, feature optimization, deep learning

1. Introduction

Aquaculture or fish farming is the application of cultivating, nurturing, and harvesting aquatic animals and weeds, including fish, shellfish, and sea plants, in subjugated environments such as ponds, lakes, and sea cages [1]. Aquaculture is a way of satisfying the increasingly wide-ranging demand for seafood and diminishing pressure on wild fish stocks. Aquaculture plays a dominant role in food security, economic establishment, and conservation of natural resources. Different aquaculture systems are used around the world, for example, freshwater and marine farming, and farmers adopt different management practices to maintain fish health and support proper growth.

The aquaculture industry makes a substantial contribution to global food production, economic growth, and employment opportunities. However, it faces several challenges, particularly those related to sustainability, environmental impact, and disease management. In 2020, worldwide aquaculture production reached

122.6 million metric tons, with a total value of approximately USD 281.5 billion [2]. Studies indicate that diseases account for almost half of the production losses in aquaculture, with the effect being more critical in developing countries. One of the primary reasons is that almost 90% of aquaculture activities are concentrated in underdeveloped nations. Fish diseases result in annual economic losses of up to USD 6 billion. Among the most commonly diagnosed fish diseases are epizootic ulcerative syndrome (EUS) (31.33%), red spot (21.33%), fin/gill rot (16.7%), grayish white spot (6%), external parasites (5%), and dropsy (4%) [3]. Aquaculture currently accounts for approximately 44% of the world's total fish production and continues to expand rapidly. Despite its remarkable growth, diseases are responsible for billions of dollars in losses each year and significantly affect production efficiency [4]. Manual detection is frequently tough because the indications are subtle and exhausting to identify, and expert knowledge is needed for accurate diagnosis, which may not always be readily accessible. Besides, poor water quality and stress could also complicate disease identification, increasing the risk of misdiagnosis and delayed treatment. This makes the transition to automated fish disease detection systems a necessity for more rapid and accurate methods.

*Corresponding author: Arslan Munir, Department of Electrical Engineering and Computer Science, Florida Atlantic University, USA. Email: arslanm@fau.edu

This research focuses on the detection of fish disease by determining whether a fish is infected or healthy. For this purpose, we collected a dataset from multiple sources containing both infected and healthy fish images. In this study, convolutional neural network (CNN) models like VGG19, VGG16, ResNet50, InceptionV3, and Xception were used for feature extraction from the mentioned dataset. The extracted features were evaluated with different machine learning (ML) classifiers, along with decision tree (DT), support vector machine (SVM), k-nearest neighbors (KNN), random forest (RF), naive Bayes (NB), extreme gradient boosting (XGB), and adaptive boosting (AdaBoost). The performance of these classifiers was tested with deep features extracted from the CNN models with and without feature optimization techniques like principal component analysis (PCA) and linear discriminant analysis (LDA). The finest method was selected by comparing the performance of all methods, which was elected as a final approach for fish disease detection. The main contributions of our work are summarized below:

- 1) Generating a dataset that includes healthy and infected fish images.
- 2) Extracting features using CNN models such as VGG19, VGG16, ResNet50, InceptionV3, and Xception and utilizing performance analysis of different classifiers with SVM, DT, KNN, RF, NB, XGB, AdaBoost, and voting ensemble classifiers for detection of fish health state.
- 3) Analyzing the outcome of feature optimization methods such as PCA and LDA for precise fish disease diagnosis.

The rest of this paper is structured as follows: Section 2 presents related work in the literature. Section 3 discusses the proposed methodology and data acquisition for this work. Section 4 presents experimental results. Section 5 discusses the results, and finally, Section 6 summarizes the study.

2. Literature Review

The automatic fish disease detection using AI assessment is centered around developing algorithms and models that can inspect fish images and detect disease symptoms [5]. In this part, we present a concise outline of recent studies related to our research objectives.

2.1. ML-based fish disease detection

Ahmed et al. [6] used image processing techniques with ML algorithms for the detection of fish disease caused by multiple pathogens, with an accuracy of 91.12% with augmentation and 94.12% without augmentation. The methodology used was image preprocessing, segmentation, feature extraction, and classification with SVM. K-means segmentation, cubic spline interpolation, and adaptive histogram equalization techniques were used to

improve the quality of the images and make them more suitable for analysis. Chakravorty et al. [7] conducted an investigation on fish mortality caused by diseases, mainly EUS, in northeast India. The study aimed to make excellent disease detection through PCA and K-means clustering using HSV images and morphological operations. However, the study identified limitations, such as a small sample size and low specificity. The authors strengthened the need for larger datasets in future studies and advised further research on real-time execution. Sikder et al. [8] proposed an automated and highly accurate approach for fish disease detection and recognition. Features were extracted from segmented regions using Gabor filters and the gray-level co-occurrence matrix (GLCM). The extracted features were then classified using M-SVM. A comparison between K-means clustering and Fuzzy C-means clustering showed that Fuzzy C-means achieved a higher accuracy of 97.90%, compared to 96.48% obtained with K-means clustering. For future work, the researchers recommended the improvement of segmentation techniques, examining more sophisticated classification algorithms and integrating deep learning (DL) techniques with variant feature descriptors. Mia et al. [9] conducted an extensive study of expert systems that can diagnose fish diseases from images taken by smartphones. They chose two different feature sets and used a segmentation technique to extract diseased regions from healthy regions of the fish. Eight widely used classification algorithms were employed and evaluated with several performance metrics. The accuracy of the most accurate classifier RF was 88.87%. The results displayed that there is scope for upliftment, and larger datasets and more fish diseases could be used for future work to raise the overall performance. Table 1 provides a synopsis of the ML-based methods.

2.2. DL-based fish disease detection

Paupi et al. [10] discussed the potential of image processing techniques to upgrade traditional approaches and the accuracy of fish disease detection. Their research had several stages of image processing, including image acquisition, preprocessing, and segmentation. The main purpose of their research was to review former studies on fish disease detection that used diverse image classification approaches such as rule-based expert systems, ML, DL, statistical approaches, and hybrid techniques. Rachman et al. [11] implemented image recognition techniques, specifically object segmentation using MobileNetV2, to find EUS in fish and gained an accuracy of 84%. This technique enhanced fish disease detection using image recognition and offered aid to farmers with limited knowledge in disease identification. Anyway, the method had difficulties detecting EUS caused by *Aphanomyces invadans*, which was identified by visual observation of red blotches on fish bodies. The authors suggested improving the model, extending it to cover additional fish diseases, and integrating advanced technologies to develop more effective fish monitoring systems.

Table 1
Existing research on fish diseases using ML methods

Research	Object	Method	Performance
Ahmed et al. [6]	Salmon fish disease	SVM	91.42%
Chakravorty et al. [7]	EUS fish disease	PCA, K-means clustering	90%
Sikder et al. [8]	Fish disease	K-means and C-means Fuzzy clustering, Gabor's filters, GLCM	96.48% (K-means), 97.90% (C-means)
Mia et al. [9]	Fish disease	Expert systems, segmentation, RF	88.87%

Krivouguz [12] investigated whether diseased fish could be accurately distinguished from healthy fish for the development of an automated real-time monitoring system for epizootic diseases in industrial aquaculture farms. In their study, a pre-trained InceptionV3 model was used to identify and classify fish based on their disease status. Two experiments were conducted, differing mainly in the use of data augmentation methods; the accuracy increased significantly to 96.9% for infected fish and 96.7% for healthy fish. These findings demonstrated the strong potential for detecting infected fish and highlighted its suitability for real-time automated fish monitoring systems. Hasan et al. [13] presented a CNN-based approach for fish disease detection. The study evaluated images of healthy fish and fish suffering from two diseases, namely, white spot and red spot. The CNN model achieved a remarkable detection accuracy of 99.44% on different test datasets. The results stated that CNNs are effective in detecting and classifying fish diseases. However, the authors mentioned that future studies with huge datasets could improve the detection performance further. Table 2 provides a summary of DL-based fish disease recognition approaches.

To overcome the limitations of existing studies, this research explores a variety of DL models and experiments with both the presence and absence of multiple feature optimization techniques. In addition, we evaluate a wide range of classifiers for fish disease recognition, whereas previous works were limited to only a small number of classifiers. Furthermore, experiments are conducted on our newly generated dataset, and the results

demonstrate that our models achieve better performance compared to existing approaches.

3. Research Methodology

Figure 1 illustrates the workflow of the proposed methodology for automated fish disease recognition, and the below sections describe this figure in detail.

3.1. Dataset

The dataset for this study contains healthy and infected fish images. The images were collected from two public repositories, namely, Kaggle and GitHub. The images were then augmented using rotation, flipping, scaling, and brightness adjustment. This results in the ultimate dataset with 1233 images of 813 healthy fish and 420 infected fish. The dataset is publicly available on GitHub, and the source is provided in [14]. Table 3 shows the distribution of fish images and some sample images from the dataset.

3.2. Image processing

Image preprocessing plays a significant role in image recognition, as it helps standardize image quality, reduce noise, and sustain relevant features. These steps improve the model’s potential to accurately classify or detect objects within the image. In

Table 2
Existing research on fish diseases using DL methods

Research	Object	DL frames	Performance
Pauzi et al. [10]	Various fish diseases	Rule-based expert system, ML, DL, statistical method, hybrid method	–
Rachman et al. [11]	Fish with EUS infection	Object segmentation, MobileNetV2	F1-score: 84%
Krivouguz et al. [12]	Fish disease	Inception v3	96.9% (for infected), 96.7% (for healthy)
Hasan et al. [13]	Healthy fish	CNN	94.44%

Figure 1
Proposed workflow diagram

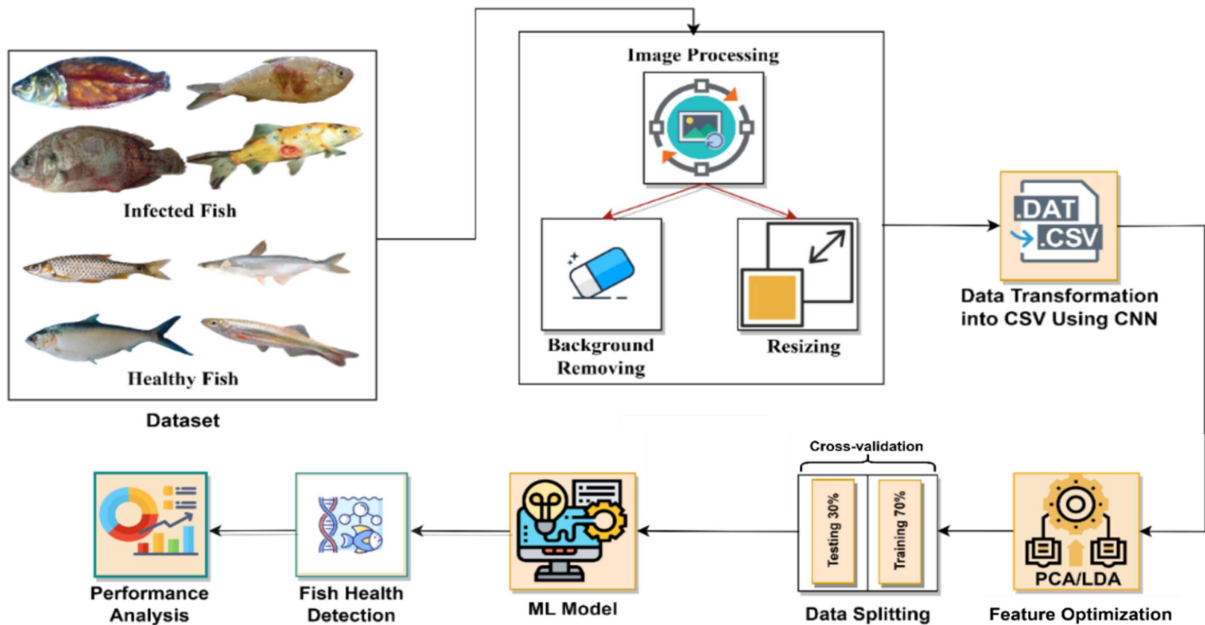


Table 3
Overview of the dataset used in this research

	Healthy Fish		Infected Fish	
Total Images	813		420	
Samples				

this study, background removal and image resizing are performed as part of the preprocessing stage before image recognition.

3.2.1. Background removal

Background removal is helpful in allowing the model to concentrate better on the initial object of interest by removing unreasonable distractions and noise. This may be a method for more suitable accuracy in object identification and classification, particularly when the images have complex or cluttered backgrounds. The method of background removal used in this study is illustrated in Figure 2.

3.2.2. Resizing

To accomplish the input size requirements of the CNN architecture used in this paper, the images are resized to 224×224 pixels, as illustrated in Figure 3. The original format of the images is resized to ensure compatibility with generally used CNN architectures such as VGG, ResNet, Inception, and Xception.

3.3. Feature extraction using CNN

In this research, we applied the pre-trained CNN models, known as VGG19, VGG16 [15], ResNet50 [16], InceptionV3 [17], and Xception [18], for feature extraction by taking advantage of their learned feature representations. The extracted features are organized and stored in a comma-separated values (CSV) file. These feature sets are then used as input for several ML classifiers to perform fish disease detection. Table 4 provides a brief comparison of the CNN architecture used in this research. Among the evaluated models, Xception is employed as the primary feature extractor in the final method. The reason for choosing it was its high capability to extract significant image patterns, which enables efficient discrimination between healthy and diseased fish. The

Xception model extracts both fine-tuned and significant features of an input image and generates detailed feature vectors.

3.4. Feature optimization

It is a widely used method in ML to minimize the number of input variables or features in a dataset. This is called feature reduction or dimensionality reduction. After feature extraction, the final CSV file will contain a large number of features, some of which can be irrelevant, redundant, or noisy. These irrelevant features can cause a higher probability of overfitting, that is, when a model performs well on training data but does not generalize well to unseen data. Two feature reduction techniques were used to solve this problem and improve the model performance, which are PCA and LDA. The techniques are discussed in detail below.

PCA: PCA is an unsupervised statistical approach, implemented to reduce the dimensionality of the data by converting the original features into a set of linearly uncorrelated elements. The modification is performed by improving the variance in the data with the covariance matrix. Let X be a dataset with n observations and p variables. The covariance matrix Σ is determined as:

$$\Sigma = \frac{1}{n-1} (X - \underline{X})^T (X - \underline{X}) \quad (1)$$

where X is an $n \times p$ matrix in which each row represents an observation, $\underline{X}$ denotes the mean vector of the variables, and T represents the transpose of a matrix. The eigenvalues and eigenvectors of the covariance matrix are then computed to identify the principal components. The eigenvectors define the directions of maximum variance in the data, while the corresponding eigenvalues indicate the amount of variance captured by each component. The original dataset is subsequently projected onto the selected principal components, resulting in a lower-dimensional representation that preserves most of the important

Figure 2
Background removal of a sample image



Before Background Removal



After Background Removal

Figure 3
Resizing of a sample image from the dataset

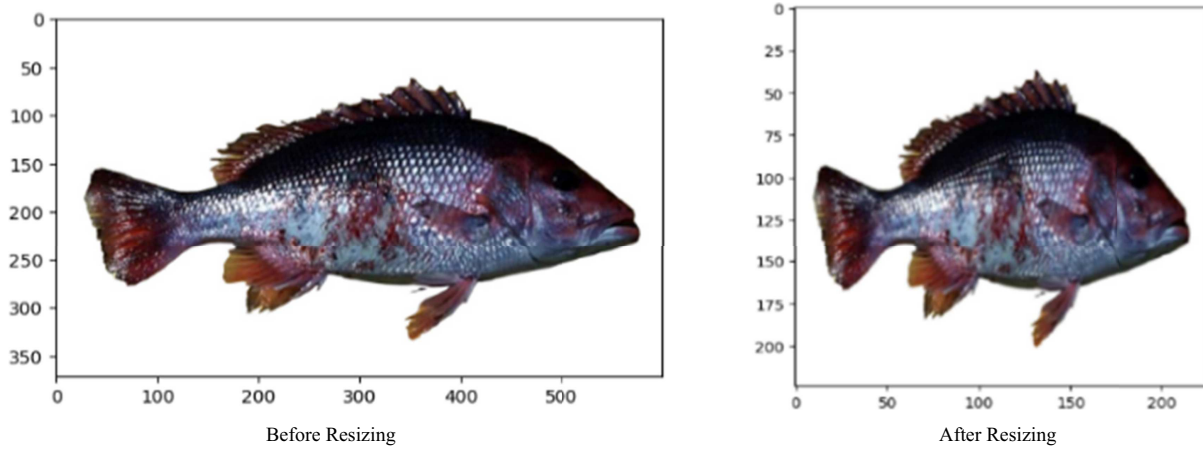


Table 4
Architectural comparison of VGG16, VGG19, ResNet50, Inception V3, and Xception models

Model	Architecture	Parameters (millions)	Features	No. of convolution layers
VGG16	Deep network with 16 weight layers	~138	4096	13
VGG19	Deep network with 19 weight layers	~143	4096	16
ResNet50	Utilizes residual learning with 50 layers	~25.6	2048	48
Inception V3	Employs Inception modules with diverse filter sizes	~23	2048	48
Xception	Features depth-wise separable convolutions	~22	2048	36

information while reducing redundancy and computational complexity [19, 20].

LDA: LDA is a supervised method of dimensionality reduction that tries to identify a feature subspace that maximizes the class separation. This is done by calculating linear discriminants, which increase the separability of the classes. The optimization procedure also involves the mathematical calculation of the within-class variance of each class, which enhances the discrimination of multiple classes.

$$\sum_{x_i \in w_j, j=1, \dots, c} W^T S_{w_j} W \quad (2)$$

where S_{w_j} denotes the within-class scatter matrix of class j , W represents the projection matrix, and c is the total number of classes. LDA determines the optimal projection directions by maximizing the ratio of between-class scatter to within-class scatter, thereby improving class separability. The transformed feature space retains the most discriminative information [21, 22].

3.5. Voting ensemble

Voting ensemble classification is an ensemble learning technique where predictions of multiple individual classifiers are combined to improve the overall performance of the model. The goal of this approach is to combine the outputs of different models, thereby reducing the prediction errors and increasing the accuracy. There are two major types of voting: hard voting, where each base classifier casts a vote for a class label and the majority vote is considered the final prediction, and soft voting, where each classifier gives probability estimates for each class and the class with the highest average probability is selected. A soft voting classifier is used in the proposed scheme. It uses the predict probability property, which returns the probability of each

target class. The training data points are passed to several classifiers, such as SVM, DT, RF, XGB, KNN, NB, and AdaBoost, after being shuffled to ensure randomness. The expected probabilities are constructed individually for each model. The final forecast is obtained by combining these probabilities using the soft voting approach and then using majority voting to choose the class with the highest combined probability [23].

Given M classifiers and C classes, let classifier i assign a probability $P_i(y = c | x)$ to class c . Let w_i denote the weight assigned to classifier i . The weighted soft voting decision is defined as:

$$\hat{y} = \arg \max_{c \in \{1, \dots, c\}} \sum_{i=0}^M \omega_i \cdot P_i(y = c | x) \quad (3)$$

Typically, the classifier weights are normalized such that $\sum_i \omega_i = 1$ [24]. Figure 4 presents the voting mechanism used in this research.

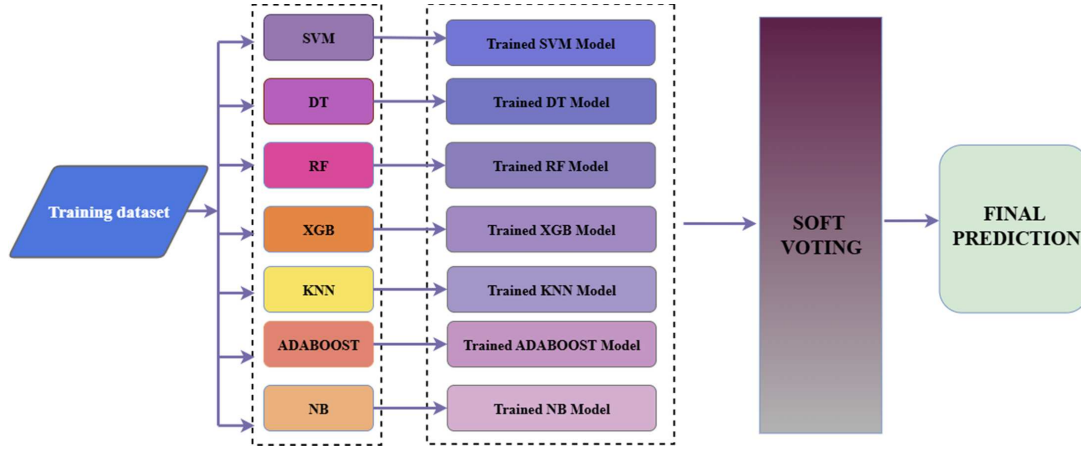
4. Results

This section presents the experimental results of our proposed methodology with the necessary criteria to measure and analyze these outcomes.

4.1. Configuration for experiments and execution

The proposed workflow is implemented using the Python 3 programming language and utilizes TensorFlow and Keras with GPU support. The entire experiment, including both training and testing phases, is conducted on the Google Colaboratory platform. The experimental environment is equipped with a Tesla T4 GPU, 12.67 GB of RAM, and 107.72 GB of available disk space. The

Figure 4
Depiction of soft voting in ensemble ML



proposed workflow is designed to classify fish images into two categories: healthy and infected.

4.2. Evaluation criteria for performance

In our research, we employ five metrics with cross-validation to assess the efficacy of the proposed model. These metrics encompass accuracy, precision, recall, F1-score, and specificity. In the following, we describe the evaluation criteria for our fish disease recognition system [25].

- 1) **Cross-validation:** Monte Carlo cross-validation (MCCV) [26] is employed to evaluate the performance of the proposed method. This research uses 10 MCCV iterations. In each cross-validation, the dataset is randomly divided into 70% training data and 30% testing data. The random partitioning process is repeated 10 times using different train-test splits. The model is trained and evaluated in each iteration. The final performance is calculated by averaging the results from all 10 runs. The repeated evaluation reduces the influence of a single random data split and provides a more reliable and robust ML model assessment.
- 2) **Accuracy:** Accuracy measures how effectively the model classifies both healthy and infected fish. It is calculated as the ratio of correctly predicted instances to the total number of predictions made. This metric provides an overall indication of the model's classification performance. The accuracy can be expressed mathematically through the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Where:

- i TN (true negative): Number of infected fish correctly classified as infected.
 - ii FP (false positive): Number of infected fish incorrectly classified as healthy.
 - iii FN (false negative): Number of healthy fish incorrectly classified as infected.
 - iv TP (true positive): Number of healthy fish correctly classified as healthy.
- 3) **Recall:** Recall measures the model's ability to correctly identify all relevant instances, specifically the proportion of correctly predicted positive cases out of all actual positive cases. In this

study, a higher recall indicates that the model is more effective in capturing most of the diseased fish cases. The mathematical expression for recall is given as

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

- 4) **Precision:** This metric measures the accuracy of healthy predictions, that is, the ratio of healthy instances to all the predicted instances as healthy. A high precision score means that when the model predicts a fish to be diseased, it is likely to be true, which reduces the false positives. The precise formulation can be defined as:

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

- 5) **F1-score:** This is the harmonic mean of precision and recall. It is a balanced measure addressing false positives and false negatives. A high F1-score in fish disease detection indicates a good balance between precision and recall. The mathematical representation of the F1-score can be articulated as follows:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

- 6) **Specificity:** This feature measures the ability of a model to correctly identify infected instances, particularly the proportion of infected instances among all actual infected instances, helping to minimize false alarms for healthy fish. The equation representing specificity can be elucidated as follows:

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

4.3. Performance analysis

Although deep CNN features can capture detailed and meaningful information, they are often high-dimensional and may contain redundant data, which can lower the performance of traditional ML classifiers. To overcome this issue, PCA and LDA were used after feature extraction to reduce unnecessary information and improve class separation. PCA helps make the features more compact by keeping the most important information, while

LDA improves the distinction between different classes. The effect of class imbalance was also resolved using the synthetic minority over-sampling technique (SMOTE) [27]. Class imbalance causes classifiers to be biased toward the majority classes and affects the reliability of classification; SMOTE solves this problem by generating synthetic minority samples by interpolating between the minority samples in the feature space. In this section, we evaluate the proposed framework using controlled ablation experiments to examine the effects of feature optimization, voting ensemble learning, and different CNN backbone models on classification performance. Three scenarios were considered: (i) no optimization, (ii) feature optimization with PCA, and (iii) feature optimization with LDA. Table 5 presents the best performer of different ablation studies in this research, and Sections 4.3.1–4.3.3 detail it.

4.3.1. Performance of different techniques without utilizing feature optimization (Ablation A)

The performance of various CNN models with different classifiers without utilizing feature optimization is evaluated. The VGG16 CNN with voting ensemble provides the highest outcome with an overall accuracy, F1-score, precision, recall, and specificity of 93.3%, 92.9%, 95.5%, 87.9%, and 96.5%, respectively. The InceptionV3 model with the KNN classifier gives the lowest outcome with an overall accuracy, F1-score, precision, recall, and specificity of 83.9%, 77.2%, 83.1%, 75.3%, and 89.2%, respectively.

4.3.2. Performance of different techniques with feature optimization PCA (Ablation B)

The performance of various CNN models with different classifiers using feature optimization PCA is evaluated. The ResNet50 CNN with the SVM classifier provides the highest outcome, with an overall accuracy, F1-score, precision, recall, and specificity of 90.2%, 89.3%, 91.5%, 90.6%, and 93.2%, respectively. The Xception model with the Voting classifier gives the lowest outcome, with an overall accuracy, F1-score, precision, recall, and specificity of 83.2%, 75.0%, 78.5%, 74.3%, and 84.8%, respectively.

With PCA only, voting and the SVM classifier provide better results with different CNNs.

4.3.3. Performance of different techniques with feature optimization LDA (Ablation C)

The performance of various CNN models with different classifiers using feature optimization LDA is evaluated. The Xception CNN with the voting classifier provides the highest outcome, with an overall accuracy, F1-score, precision, recall, and specificity of 98.3%, 97.7%, 98.6%, 97.1%, and 99.1%, respectively. The InceptionV3 model with the voting classifier gives the lowest outcome, with an overall accuracy, F1-score, precision, recall, and specificity of 94.4%, 91.3%, 95.5%, 89.5%, and 97.3%, respectively. With LDA, all CNNs perform well with the voting classifier, except ResNet50, which works well with XGB. Since this study addresses a binary classification problem, LDA generates a single discriminant component (output dimension = 1). The singular value decomposition (SVD) solver was used without shrinkage, and in each MCCV iteration, LDA was fitted on the training data only. The test data were transformed using the LDA projection learned from the corresponding training set without being used to estimate the transformation.

4.3.4. Key performance insights

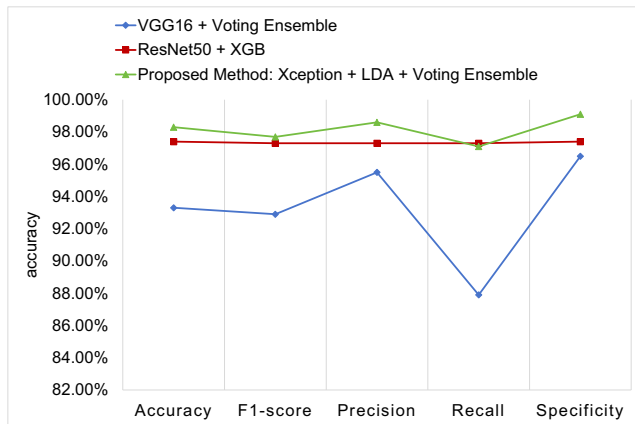
Figure 5 presents a performance comparison of the three best-performing models. The comparison of the performance of VGG16 + Voting Ensemble, ResNet50 + XGB, and the proposed method (Xception + LDA + Voting Ensemble) for five metrics. The proposed method consistently outperforms the baselines with accuracy, precision, and specificity values above 98%. ResNet50 + XGB consistently performs around 97%, whereas VGG16 + Voting Ensemble provides comparatively poorer results, especially in recall. In conclusion, the figure shows the improved robustness of the proposed method.

To validate the efficiency of the final model, a detailed evaluation was performed using a confusion matrix and receiver operating characteristic (ROC) curve, as illustrated in Figures 6(a) and (b). From the confusion matrix, it can be seen that the model correctly classified 97.1% infected samples and 99.1% healthy

Table 5
Summary of the best performer of different ablation studies in this research

Study	Feature optimization	CNN	Best classifier	Accuracy	F1-score	Precision	Recall	Specificity
Ablation A	None	VGG19	SVM	89.8	87.2	90.6	89.8	94.1
		VGG16	Voting	93.3	92.9	95.5	87.9	96.5
		ResNet50	SVM	90.9	87.9	91.7	90.9	93.2
		Xception	XGB	87.0	82.0	85.5	81.5	90.7
		InceptionV3	KNN	83.9	77.2	83.1	75.3	89.2
Ablation B	PCA	VGG19	SVM	88.2	87.2	89.6	89.8	90.1
		VGG16	SVM	89.7	88.5	90.9	89.5	91.3
		ResNet50	SVM	90.2	89.3	91.5	90.6	93.2
		Xception	Voting	83.2	75.0	78.5	74.3	84.8
		InceptionV3	Voting	83.7	77.0	78.4	73.1	88.4
Ablation C	LDA	VGG19	Voting	97.2	96.3	95.9	97.1	97.3
		VGG16	Voting	95.5	93.6	95.9	92.4	97.3
		ResNet50	XGB	97.4	97.3	97.3	97.3	97.4
		Xception	Voting	98.3	97.7	98.6	97.1	99.1
		InceptionV3	Voting	94.4	91.3	95.5	89.5	97.3

Figure 5
Comparative performance of baseline models and the proposed method across evaluation metrics



samples. Only 2.9% infected samples were misclassified as healthy, and 0.9% healthy samples were misclassified as infected. The performance results showed high classification ability and low misclassification rate, which are critical in sensitive areas like fish disease diagnosis, where accurate identification is very important. Furthermore, the ROC curve was employed to evaluate the model's discrimination capability across varying thresholds between classes. The area under the curve (AUC) was found to be 0.982, indicating excellent classification performance. The high value of AUC of the ensemble model close to 1.0 validates high levels of separability of the positive class and the negative class, which proves its robustness and generalization ability. The near-perfect shape of the ROC curve and high AUC score validate the effectiveness of an ensemble of multiple classifiers to build a highly reliable and accurate prediction system for automated fish disease detection.

5. Discussion

This study explores the effectiveness of optimized deep features for automatic fish disease detection by combining different CNN models with traditional ML classifiers. The research

demonstrates the performance improvement and also explains the effect of feature optimization and ensemble learning on the classification results. Deep CNN models can extract detailed and high-level features. However, too many features sometimes can bring redundant or unnecessary data, which may degrade the performance of traditional classifiers. Ablation experiments show that PCA, an unsupervised reduction technique in dimension, provides minor advantages and mainly benefits classifiers like SVM. Despite this, LDA consistently boosts the performance of various CNN models, especially deeper models such as Xception and ResNet50. This indicates that LDA does not reduce the deep features but makes the differentiation among different classes clearer and improves the classification accuracy by supervised feature projection. Furthermore, the voting ensemble method improves the results by compensating for the weaknesses of individual classifiers, making the overall system more robust and stable. The proposed framework provides a more comprehensive evaluation of CNN-based fish disease detection compared to related work. Most of the existing studies are limited to a single CNN architecture or a limited number of classifiers, which reduces their applicability across different disease categories. Segmentation-based methods using MobileNetV2 [11] or classification systems that use only InceptionV3 [12] get good results but are architecturally limited. Other methods combine handcrafted features or classical image processing pipelines with single classifiers [6, 7, 9]. As a result, this work provides a systematic evaluation of several CNN backbones and classifiers under controlled ablation conditions and shows that optimized deep features together with ensemble learning can lead to consistent and scalable performance improvements.

Table 6 summarizes representative state-of-the-art methods alongside the proposed approach. While several existing models report competitive accuracy, they often lack explicit feature optimization or ensemble strategies. The proposed Xception + LDA + Voting Ensemble configuration achieves the highest accuracy of 98.3%, underscoring the practical benefit of supervised feature optimization when applied to deep CNN representations. To enhance readability and emphasize scientific insight, detailed classifier-backbone performance results are provided in the supplementary material, while the main text focuses on key trends and the best-performing configurations. Despite these encouraging results, several limitations remain. The dataset size,

Figure 6
Confusion matrix (a) and ROC curve (b) of the selected ensemble model (Xception + Voting Classifier + LDA)

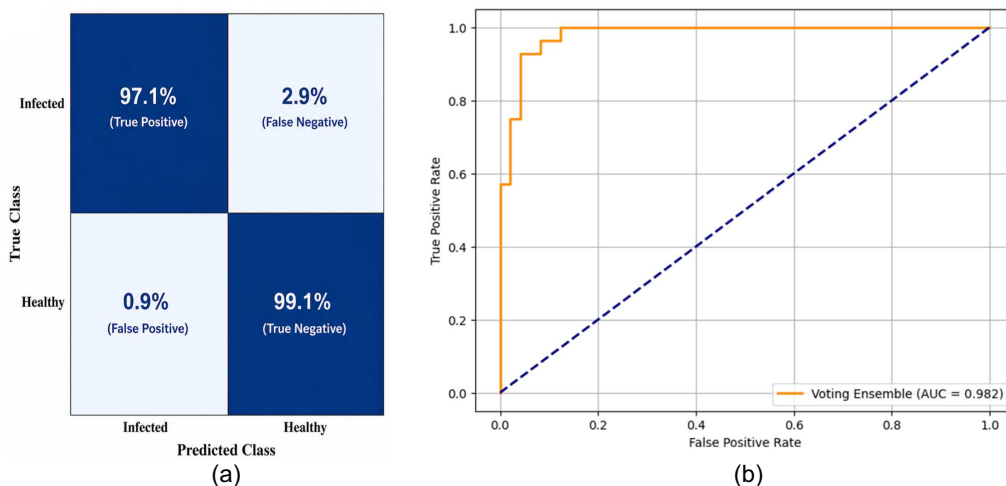


Table 6
Evaluate the proposed model in comparison to previously published state-of-the-art models

Study	Architecture	Classifier	Accuracy (%)
Rachman et al. [11]	MobileNetV2	–	84
Krivoguz et al. [12]	Inception v3	–	86.8
Hasan et al. [13]	CNN	–	94.44
Chakravorty et al. [7]	PCA-based recognition	K-means clustering	90
Ahmed et al. [6]	–	SVM	94.12
Mia et al. [9]	Segmentation algorithm	RF	88.87
Proposed Method	Xception and LDA	Voting	98.3

although comparable to or larger than those used in several related studies, is still moderate; future work will therefore focus on expanding the dataset to include a broader range of fish species and disease conditions. In addition, hyperparameter tuning in the current study was conducted within computationally feasible ranges guided by prior literature; however, exploring wider search spaces and more advanced optimization strategies, such as Bayesian optimization, may further enhance model performance. Future investigations will also consider validation on independent datasets and the adoption of end-to-end fine-tuning strategies to improve generalization and applicability in real-world aquaculture environments.

6. Conclusion

Our study highlights the importance of computer-aided diagnosis (CAD) systems to address infectious diseases in the aquaculture industry. This study aims to improve the detection of fish diseases by combining DL architectures such as VGG16, VGG19, ResNet50, InceptionV3, Xception, and multiple classifiers. Techniques for feature optimization, such as PCA and LDA, were also included to improve the accuracy of disease identification. The proposed workflow illustrates the importance of the use of multiple data sources, image processing techniques, and a structured dataset of 1233 fish images. It achieved promising results in fish disease classification. For reliable evaluation, cross-validation was implemented. This provided a detailed analysis of performance metrics like accuracy, F1-score, precision, recall, and specificity for various combinations of models and classifiers. Further studies with feature optimization techniques have shown their effectiveness in enhancing model performance and have helped to better understand the strengths and limitations of each combination. The implementation of CAD systems in aquaculture could help reduce economic losses, support environmentally sustainable practices, and strengthen the seafood industry against globally spreading diseases. In future work, we plan to develop electronic tools for real-time dish health monitoring. We also aim to explore more recent DL and ML techniques to further improve disease identification. In addition, our image dataset will be expanded by including more images of both healthy and infected fish to support the development of more accurate diagnostic models.

Ethical Statement

The authors declare that this study did not require formal ethical approval because Bangladesh University does not require Institutional Review Board or ethics committee approval for

research that uses secondary data available online. This exemption is based on the research integrity policy, issued by the Department of CSE, Bangladesh University.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in GitHub at <https://github.com/meharaj-emon028/Fish-Dataset>, reference number [14].

Author Contribution Statement

Md. Meharaj Uddin: Methodology, Data curation, Writing – original draft. **Ummatun Nahar Oishi:** Methodology, Data curation, Writing – original draft. **Arslan Munir:** Conceptualization, Validation, Investigation, Resources, Writing – review & editing, Supervision. **Muhammad Minoar Hossain:** Conceptualization, Formal analysis, Writing – review & editing, Project administration. **Md. Khabir Uddin Ahamed:** Software, Investigation, Visualization. **Md. Sadiq Iqbal:** Software, Resources.

References

- [1] Vij, S., Lee, Y. S., Purushothaman, K., & Jerry, D. (2024). Challenges in Singapore aquaculture and possible solutions. *Aquaculture Journal*, 4(4), 316–323. <https://doi.org/10.3390/aquacj4040023>
- [2] Food and Agriculture Organization of the United Nations. (2022). *FAO report: Global fisheries and aquaculture production reaches a new record high*. <https://www.fao.org/newsroom/detail/fao-report-global-fisheries-and-aquaculture-production-reaches-a-new-record-high/en>
- [3] Bagum, N., Monir, M. S., & Khan, M. H. (2013). Present status of fish diseases and economic losses due to incidence of disease in rural freshwater aquaculture of Bangladesh. *Journal of Innovation & Development Strategy*, 7(3), 48–53.
- [4] Assefa, A., & Abunna, F. (2018). Maintenance of fish health in aquaculture: Review of epidemiological approaches for prevention and control of infectious disease of fish. *Veterinary Medicine International*, 2018(1), 5432497. <https://doi.org/10.1155/2018/5432497>

- [5] Li, D., Li, X., Wang, Q., & Hao, Y. (2022). Advanced techniques for the intelligent diagnosis of fish diseases: A review. *Animals*, 12(21), 2938. <https://doi.org/10.3390/ani12212938>
- [6] Ahmed, M. S., Aurpa, T. T., & Azad, Md. A. K. (2022). Fish disease detection using image based machine learning technique in aquaculture. *Journal of King Saud University—Computer and Information Sciences*, 34(8), 5170–5182. <https://doi.org/10.1016/j.jksuci.2021.05.003>
- [7] Chakravorty, H., Paul, R., & Das, P. (2015). Image processing technique to detect fish disease. *International Journal of Computer Science and Security*, 9(2), 121–131.
- [8] Sikder, J., Sarek, K. I., & Das, U. K. (2021). Fish disease detection system: A case study of freshwater fishes of Bangladesh. *International Journal of Advanced Computer Science and Applications*, 12(6), 867–871. <https://doi.org/10.14569/IJACSA.2021.01206100>
- [9] Mia, M. J., Mahmud, R. B., Sadad, M. S., Asad, H. A., & Hossain, R. (2022). An in-depth automated approach for fish disease recognition. *Journal of King Saud University - Computer and Information Sciences*, 34(9), 7174–7183. <https://doi.org/10.1016/j.jksuci.2022.02.023>
- [10] Pauzi, S. N., Hassan, M. G., Yusoff, N., Harun, N. H., Abu Bakar, A. H., & Kua, B. C. (2021). A review on image processing for fish disease detection. *Journal of Physics: Conference Series*, 1997(1), 012042. <https://doi.org/10.1088/1742-6596/1997/1/012042>
- [11] Rachman, F., Akbar, M. N. S., & Putera, E. (2023). Fish disease detection of Epizootic Ulcerative Syndrome using deep learning image processing technique. In *Proceedings of the 9th International Conference on Fisheries and Aquaculture*, 8(1), 23–34. <https://doi.org/10.17501/23861282.2023.8102>
- [12] Krivoguz, D. (2023). Deep learning approach of using Inception deep neural network for real-time epizootic monitoring of fish diseases. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.12502.93761>
- [13] Hasan, N., Ibrahim, S., & Aqilah Azlan, A. (2022). Fish diseases detection using convolutional neural network (CNN). *International Journal of Nonlinear Analysis and Applications*, 13(1), 1977–1984. <https://doi.org/10.22075/ijnaa.2022.5839>
- [14] Meharaj. (2024). *Fish-dataset [Data set]*. *GitHub*. <https://github.com/meharaj-emon028/Fish-Dataset>
- [15] Ahamed, K. U., Islam, M., Uddin, A., Akhter, A., Paul, B. K., Yousuf, M. A., . . . , & Moni, M. A. (2021). A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images. *Computers in Biology and Medicine*, 139, 105014. <https://doi.org/10.1016/j.compbiomed.2021.105014>
- [16] Wan, J., Li, B., Wang, K., Teng, X., Wang, T., & Mao, B. (2024). An improved ResNet50 for environment image classification. *Procedia Computer Science*, 242, 1000–1007. <https://doi.org/10.1016/j.procs.2024.08.246>
- [17] Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Díez, I. D. L. T., & Ashraf, I. (2022). Pneumonia classification from X-ray images with Inception-V3 and convolutional neural network. *Diagnostics*, 12(5), 1280. <https://doi.org/10.3390/diagnostics12051280>
- [18] Kassani, S. H., Kassani, P. H., Khazaeinezhad, R., Wesolowski, M. J., Schneider, K. A., & Deters, R. (2019). Diabetic retinopathy classification using a modified Xception architecture. In *2019 IEEE International Symposium on Signal Processing and Information Technology*, 1–6. <https://doi.org/10.1109/ISSPIT47144.2019.9001846>
- [19] Carobene, A., Campagner, A., Uccheddu, C., Banfi, G., Vidali, M., & Cabitza, F. (2022). The multicenter European Biological Variation Study (EuBIVAS): A new glance provided by the Principal Component Analysis (PCA), a machine learning unsupervised algorithms, based on the basic metabolic panel linked measurands. *Clinical Chemistry and Laboratory Medicine*, 60(4), 556–568. <https://doi.org/10.1515/cclm-2021-0599>
- [20] Sgarro, G. A., Vasco, P., Santoro, D., Grilli, L., Giglio, M., Brunetti, N. D., . . . , & Valenzano, A. A. (2026). Identifying risk patterns for sudden cardiac death in athletes: A clustering and principal component analysis approach. *PLOS One*, 21(1), e0339377. <https://doi.org/10.1371/journal.pone.0339377>
- [21] Zhao, S., Zhang, B., Yang, J., Zhou, J., & Xu, Y. (2024). Linear discriminant analysis. *Nature Reviews Methods Primers*, 4(1), 70. <https://doi.org/10.1038/s43586-024-00346-y>
- [22] Deng, T., Huang, J., & Chen, Y. (2026). Metric learning based weighted linear discriminant analysis for imbalanced multi-label classification. *Applied Intelligence*, 56(1), 30. <https://doi.org/10.1007/s10489-025-07045-5>
- [23] Kumari, S., Kumar, D., & Mittal, M. (2021). An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *International Journal of Cognitive Computing in Engineering*, 2, 40–46. <https://doi.org/10.1016/j.ijcce.2021.01.001>
- [24] Manconi, A., Armano, G., Gnocchi, M., & Milanese, L. (2022). A soft-voting ensemble classifier for detecting patients affected by COVID-19. *Applied Sciences*, 12(15), 7554. <https://doi.org/10.3390/app12157554>
- [25] Hossain, M. M., Swarna, R. A., Mostafiz, R., Shaha, P., Pinky, L. Y., Rahman, M. M., . . . , & Iqbal, M. S. (2022). Analysis of the performance of feature optimization techniques for the diagnosis of machine learning-based chronic kidney disease. *Machine Learning with Applications*, 9, 100330. <https://doi.org/10.1016/j.mlwa.2022.100330>
- [26] Dias, M. J. G., de Jesus Torres, L., de Oliveira Silva, H., Harmon, J. R., Júnior, R. G. S., & de Oliveira Nascimento, F. A. (2025). Statistical evaluation of machine learning classifiers using Monte Carlo Cross-Validation for breast cancer diagnosis. In *2025 Workshop on Communication Networks and Power Systems*, 1–7. <https://doi.org/10.1109/WCNPS69127.2025.11295901>
- [27] Khorshidi, H. A., & Aickelin, U. (2025). A synthetic over-sampling method with minority and majority classes for imbalance problems. *Knowledge and Information Systems*, 67(7), 5965–5998. <https://doi.org/10.1007/s10115-025-02394-6>

How to Cite: Uddin, M. M., Oishi, U. N., Munir, A., Hossain, M. M., Uddin Ahamed, M. K., & Iqbal, M. S. (2026). An Intelligent Scheme of Automatic Fish Disease Recognition Using Optimized Deep Features. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA62025761>