Journal of Advanced Information and Intelligent Technology, Volume 21, Issue 2, July 2026

Received: 13 May 2025 | Revised: 23 June 2026 | Accepted: 25 June 2026

DOI: https://doi.org/10.36481/jaiit.v21ino2.9xvb2375

RESEARCH ARTICLE

Publisher: Daffodil International University
https://journals.diu.edu.bd/index.php/jaiit

Journal of Advanced Information and Intelligent Technology

Fresh Fish Disease Detection in Aquaculture: Leveraging CNN Models for Accurate Diagnosis

*Tapasy Rabeya[1], Sonia Nasrin[2], Lamia Rukhsara[1], Israt Jahan[1] and Umme Habiba[1]
[1]Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
[2]Department of Computing and Information System, Daffodil International University, Dhaka, Bangladesh

*Email: tapasyrabeya.cse@diu.edu.bd
ORCID iD: 0000-0003-3111-2489

Abstract: Aquatic products such as cultivating freshwater fish are indispensable for meeting the world's protein demands. It becomes imperative to raise the fish production. But the industry faces a significant obstacle in the form of disease outbreaks among aquatic populations. These outbreaks not only lead to significant financial losses but also give rise to severe ecological concerns. The identification of diseased fish in aquaculture is still difficult due to lack of necessary infrastructure. It is crucial to identify contaminated fish at an early stage to stop the disease from spreading. Thus, a study on fish disease detection using CNN is proposed. A sample of about 3,000 fish images was gathered from open-source repositories with seven classes, including six typical diseases of freshwater fish and healthy fish. Resizing, noise reduction, normalization, and data augmentation of the images were performed beforehand to improve the generalization of the model. Multi-class disease classification VGG16, VGG19, InceptionV3, DenseNet, MobileNet and ViT, are six pretrained CNN architectures that were implemented and tested. Experimental findings show that DenseNet and MobileNet were the most precise, with the greatest classification accuracy of 90% and the highest recall as well as F1-score. This suggests that the CNN performs a reliable job of identifying and categorizing the type of diseases among infected fishes and provide a dependable and scalable solution to automated disease surveillance in aquaculture Furthermore, its focus on real-world image conditions, including noise and quality variations in photographs captured directly from aquaculture environments, makes it more suitable for practical deployment and automated field-level disease monitoring.

Keywords: *Image Processing, Deep Learning, Artificial Intelligence, Computer Vision, Fisheries.*

1. Introduction

Aquaculture is a process of raising aquatic creatures under regulated circumstances, including fish, crustaceans, mollusks, and aquatic plants as shown in figure 1. It entails the breeding, growing, and harvesting of these creatures in freshwater, brackish water, and saltwater habitats [1]. In a world where over 70% of the earth is covered in water and about 30% of people suffer from malnutrition, aquaculture is a key contributor to the world food supply, particularly in light of the rising need for sustainable protein sources [2]. Aquaculture is a crucial option to offer food security and satisfy nutritional demands as the human population expands and conventional fisheries suffer depletion.

Fig.1. Aquaculture of fresh water in rural area [3]

Aquaculture may be classified into three broad categories. The former is the culture of the sea, or, in other words, the cultivation of marine life, both flora and fauna. Various technologies are used in the world to make marine products [4]. The culture of the Mari involves the cultivation of organisms as a source of food, non-food (e.g., pharmaceuticals, food additives, jewelry (e.g., cultured pearls), nutraceuticals), and cosmetics. These can be practiced in the natural marine setting or in confined habitats like cages, ponds, and raceways, either on land or on the sea. Some of the common species that are farmed are seaweeds, mollusks, shrimp, marine fish, and other organisms like sea cucumbers and seahorses [5]. According to the current population growth and consumption trends, aquatic food production will be 83 million tons in the year 2030, which will be an augmentation of 37.5 million tons from the year 2004. Currently, the mariculture contributes about 51 percent of the total aquatic production, indicating that over 42 million tons of aquatic products are likely to be produced in the future around the world, which will pose great difficulties in satisfying the global demand [5]. The second type is that of fish farming, which deals with the cultivation of fish in artificial habitats like ponds, tanks, and cages as a commercial enterprise.

Fish farming has grown at a high rate in recent years to satisfy the growing demand of fish products. Popular species that are grown are catfish, tilapia, salmon, carp, cod, and trout species [6]. The third type is brackish-water aquaculture, where living organisms grow in water that contains a salinity that lies between freshwater and seawater. With the rising sea levels, the brackish-water aquaculture is likely to expand further, providing seaweed, shellfish, and finfish, mainly to be used by humans. Nevertheless, its aggressive exploitation of natural resources and its production of waste may have both direct and indirect impacts on biodiversity. Since the water conditions and tidal energy are favorable, mangrove ecosystems have been the main recipients of brackish-water aquaculture, especially in shrimp farming [7].

The importance of aquatic products in the world food supply chain has been growing and has had a tremendous economic and social effect, especially in the developing nations [8]. The consumption of fish is on the increase since fish is one of the main sources of high-quality protein as well as a rich source of the balanced combination of the necessary nutrients needed by the human body [9]. Nevertheless, the high stocking density and intensification of aquaculture systems have raised serious questions about fish health. Research has shown that about 50 percent of the overall production losses are caused by fish diseases. Repeat disease outbreaks and fast spreading may cause huge numbers of people to be infected in a short duration, causing high mortality and water pollution [10]. In addition, fish that are contaminated with pathogenic microorganisms may also be hazardous to human health, including the infection of Salmonella [11]. Late detection of the disease might cause the total extinction of farmed fish. Consequently, to ensure fish welfare, as well as sustainable aquaculture practices, the creation of non-destructive, real-time, rapid, and automated disease detection and prediction systems is necessary [8].

The image processing techniques have been extensively used to cut manual labor and processing time without deteriorating the accuracy of classification. The methods can be widely used in the field of fish species identification, which is essential to track the ecological behavior and estimate fish stock levels and consumption patterns [12]. However, it is not possible to fully eradicate human control based on the lack of scientific evidence. In the recent past, machine learning and deep learning algorithms have shown great promise of accurate fish using image information, which greatly decreases the necessity of human intervention [13, 14]. Inspired by such developments, this paper targets the classification of freshwater fish diseases using images. The suggested framework groups fish images into 7 categories, namely: Bacterial Red Disease, Aeromoniasis, Bacterial Gill Disease, Parasitic Diseases, Viral Diseases, and Healthy Fish. It uses five major deep learning models, including VGG16, VGG19, InceptionV3, DenseNet, MobileNet and ViT to extract discriminative features and accurately classify infectious fish diseases.

2. Literature Review

To overcome the challenges listed in introduction section, a system with feature variability is required. Therefore, there are many studies that have investigated fish classification and disease detection based on data mining, machine learning, and deep learning methods. Muhammad et al. developed a deep convolutional neural network (CNN)-based fish species recognition model with a dropout layer, as the selection of a dropout layer enhanced the classification accuracy of the model by 86.65 percent to 90.48 percent [15]. Likewise, Kristian et al. proposed a deep learning architecture that effectively examines a large amount of image data without any filtering. Their method first identifies fish in single images before using a CNN that has a squeeze-and-excitation architecture. The ImageNet pretrained model attained 99.27% and 87.74% pre-training and post-training accuracy, respectively. Furthering these methods, Ahmed et al. examined the detection of salmon fish disease in aquaculture based on a recently developed dataset. Their experiment compared the performance of classification with and without image augmentation. The model also hit accuracy rates of 91.42 and 94.12 with and without data augmentation after using a Support Vector Machine (SVM) classifier, respectively [1]. Ahsan et al. created a YOLO-based fish detection and population density estimation system in another study that combines the use of optical flow, Gaussian mixture modeling, and deep learning. Their approach presented accuracies of 91.64 and 79.8 percent on two different datasets, which showed that it was strong when used in different environmental conditions [16].

However, in contrast to deep learning approaches, Malik et al. performed experiments in a MATLAB environment on actual images of an Epizootic Ulcerative Syndrome (EUS)-infected fish dataset. Aphanomyces invadans is a fungal pathogen that causes EUS. The authors used the Histogram of Oriented Gradients (HOG) and Features from Accelerated Segment Test (FAST) to extract features and proved that the feature dimensionality in the case of Principal Component Analysis (PCA) yielded much better classification accuracy [17]. In the same line of activity, Sikder et al. introduced an automated system of freshwater fish disease detection by using K-means and C-means fuzzy and clustering systems to image segmentation, in which fish samples were taken in the Rangamati Kaptai Lake and Sunamganj Hoar areas of Bangladesh. The feature extraction was done by the use of Gabor filters and Gray Level Co-occurrence Matrix (GLCM) methods, and classification was done with the Multi-Support Vector Machines (M-SVM) [18].

Bacterial and viral pathogens are commonly known to cause fish infections and they can substantially change the most important water quality parameters, pH, dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), total dissolved solids (TDS), electrical conductivity (EC), phosphate (PO 4 -), nitrate (NO 3 -N Based on this association, Nayan et al. concentrated on the early warning and forecasting of water quality loss using machine learning to avert the outbreak of diseases in aquaculture systems [19]. In addition, Desai et al. performed a comparative analysis of the feature extraction ability with state-of-the-art deep learning networks, such as CNN, VGG19, Inception, Xception, ResNet, and EfficientNet. Their experimental findings showed that models of image processing using deep learning gave better performance with the highest classification accuracy of 100%. Besides this, Sebastian and co-authors made a comparison between CNN-based fish identification and human identification and revealed that the accuracy of CNN models was 94.9 percent, which is significantly higher than the accuracy of human identification, which is 84.9 percent [20]. Lastly, Alaa et al. suggested an identity framework of fish based on biometrics with a success rate of 96.4%. They used a feature extraction and dimensionality reduction with Linear Discriminant Analysis (LDA) and an eventual classification with an AdaBoost classifier [21].

The past ten years have gone by with the fast pace of using deep learning methods to monitor the health of aquaculture. Recent systematic reviews point out that image analysis with a convolutional neural network (CNN) can yield strong feature extraction and precise classification ability of a diverse array of fish diseases, which is significantly better than conventional handcrafted feature-based methods [22]. Such works also highlight the usefulness of transfer learning and lightweight models, as well as MobileNet, and densely connected networks, including DenseNet, that are both highly diagnostic and resource-efficient to apply to aquaculture farms with resource constraints in real-time and edge devices [22, 23]. Simultaneously, hybrid detection systems, combining temporal or motion data with deep learning systems, have been shown to be highly effective in unconstrained underwater settings, especially in adverse lighting and turbidity situations, suggesting they should be considered in disease surveillance in video streams [24]. Recent applied studies also state that object-detection pipelines that utilize versions of YOLO with or without attention mechanisms can identify diseased areas on fish bodies in near real time, enabling automated lesion tracking and early warning systems in intensive aquaculture [25]. All in all, the mentioned advances highlight the increased role of automated, vision-based disease detection systems to minimize the reliance on expert supervision, promptly intervene, and reduce economic losses caused by diseases in the aquaculture sector.

Previous studies showed good performance in fish classification and disease detection using different machine learning and deep learning methods. However, many of those studies focused on limited disease categories or specific models. In this study, five pretrained CNN models were compared using a multi-class freshwater fish disease dataset. Among them, DenseNet and MobileNet achieved the highest classification accuracy of 90%, which indicates that the proposed framework is effective and reliable for freshwater fish disease detection in aquaculture systems.

3. Methodology

The proposed workflow covers the entire process, beginning with dataset construction and proceeding through data preprocessing, model implementation, and final disease classification. Dataset preparation was carried out prior to model training, as the availability of high-quality labeled data is a critical factor in the performance of machine learning models. After applying appropriate preprocessing and augmentation techniques, multiple deep learning–based classification models were implemented and evaluated. The overall procedure is divided into several stages, which are explained in detail through subsections A to C in this methodology section. The following section outlines the multi-stage approach that will be used in this research, as shown in Fig. 2. A logical and organized research methodology was adhered to in order to have a correct and dependable classification of freshwater fish diseases.

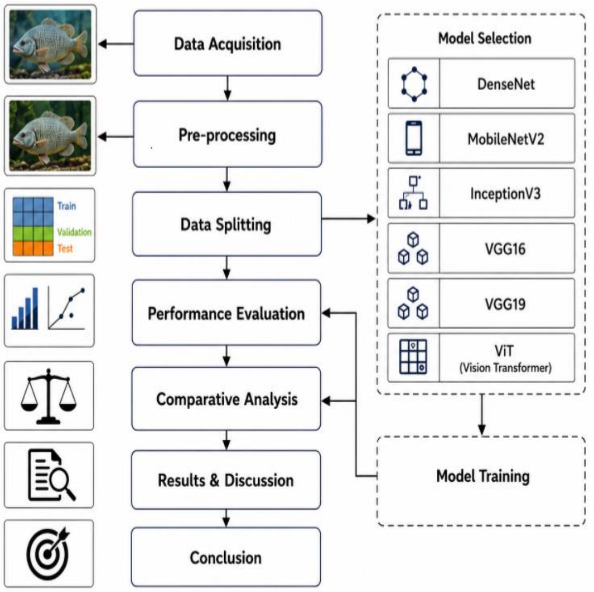


Fig.2. The working methodology of the model

3.1. Data Description

This study utilizes image data collected from publicly available open-source repositories. Specifically, the dataset was compiled from open-access online sources, including Kaggle and other publicly shared academic image repositories that provide fish disease-related images for research purposes. All images used in this study are freely accessible, do not contain any personal or sensitive information, and are ethically suitable for academic research. Prior to model training, the collected images were manually reviewed to ensure correct labeling, sufficient visual quality, and relevance to freshwater aquaculture disease classification. The last dataset is made up of pictures of seven different classes, with six classes of diseased fish and one class of healthy fish. Aquaculture diseases in fish may seriously influence growth, survival, and general productivity, and are usually due to bacteria, viruses, parasites, fungi, and environmental agents. In accordance, this study on the classification of the following categories: Bacterial Red Disease, Aeromoniasis, Bacterial Gill Disease, Fungal Disease (Saprolegniasis), Parasitic Diseases, Viral Diseases, and Healthy Fish are highlighted in fig. 3 and fig.4.

The bacteria Pseudomonas fluorescens are the primary cause of the bacterial Red Disease (BRD). The disease is marked by bleeding and skin, fin, and gill ulcers in the affected fish. The typical clinical symptoms are skin redness, ulceration, anorexia, and lethargy [26].
Aeromoniasis or Motile Aeromonas Septicemia (MAS) is an infectious disease caused by bacteria belonging to the genus Aeromonas (and in particular Aeromonas hydrophila). It impacts extensive groups of freshwater and marine fish. The most common ones are hemorrhagic septicemia, skin ulcers, exophthalmia (protruding eyes), and ascites, which are the presence of fluid in the abdominal cavity [27].

Flavobacterium branchiophilum is the main causative agent of Bacterial Gill Disease and attacks the gill tissues mainly. The symptoms of infected fish are swollen and discolored gills, excessive secretion of mucus, breathing difficulty (gasp at the surface of the water), and lethargy [28].

Saprolegniasis is a fungal infection through organisms of the genus Saprolegnia. It commonly affects fish eggs and fish with external injuries. The disease is characterized by cotton-like growths on the skin, gills, or eggs and may be accompanied by secondary bacterial infections, reduced appetite, and fatigue [29].

Parasitic diseases in fish are caused by various organisms, including protozoans, helminths, and crustaceans. Among the most prevalent parasitic infections are *Ichthyophthirius multifiliis* (commonly known as white spot disease) and *Gyrodactylus* species (skin flukes). Infected fish may display white spots on the skin and gills, excessive mucus production, flashing behavior (rubbing against surfaces), emaciation, and lethargy [30].

Viral diseases are among the most severe infections affecting fish and often result in high mortality rates. Notable examples include Infectious Hematopoietic Necrosis Virus (IHNV) and Viral Hemorrhagic Septicemia Virus (VHSV). Clinical symptoms of viral infections include hemorrhages in the skin and internal organs, skin darkening, exophthalmia, and abdominal distention [31]. In fig.3 the affected fish are shown:

(a) Bacterial Red disease

b) Acterial diseases – Aeromoniasis

(c) Bacterial gill disease

(d) Fungal diseases Saprolegniasis

(e) Parasitic diseases

(f) Viral diseases White tail

Fig. 3. Sample data of disease affected fresh water fish.

Aquaculture systems require the production of high productivity and long-term sustainability of healthy fish. Healthy fish have normal behavior, healthy growth rates, and no signs of disease or physiological stress. The main aspects of fish health, such as balanced nutrition and proper management of water quality, optimum stocking ratios, as well as disease prevention and control, are the main aspects of fish health [32]. The sample data of healthy fish are presented in fig.4.

Fig. 4. Healthy fish sample of the dataset.

3.2. *Data Preprocessing*

The image augmentation was used to deal with the issue of class imbalance and improve the generalization ability of the proposed models. In order to enhance the quality of data and to facilitate its compatibility with deep learning architectures, a number of image preprocessing methods were used, as follows.

- ***Image Resizing:*** The raw images in the dataset were different in terms of resolution and dimensions. Thus, all images were down sampled to 224 x 224 pixels, a typical convolutional neural network input, especially for pretrained models like VGG16 and ResNet.
- ***Noise Reduction:*** To minimize noise in images, a Gaussian filtering method was used, but the filter did not destroy critical visual information that is associated with fish diseases. Being an isotropic low-pass filter, the Gaussian filter averages out high-frequency noise and thus smooth images, but does not cause a serious loss to significant structural information.
- ***Augmentation and Normalization:*** To enhance the robustness of the model and avoid its overfitting, the data augmentation methods were implemented. Each picture was scaled to the range 0-1 by dividing by the maximum possible pixel value of 255 to ensure the numbers are stable in training. Horizontal flipping, width and height shifting (0.2), shearing (0.2), and zooming (0.2) were used as augmentation operations. Such a transformations model real-world changes in image orientation and scale, and hence enhances the model to extrapolate to unseen data.

- ***Dataset Splitting:*** The whole dataset is divided into 3,000 images. Among them, 2,398 images were utilized as training, 300 as validation, and 299 as testing. This method of data partitioning guarantees sound performance assessment and minimizes the possibility of overfitting.

3.3. *Model Discussion*

In this study, multiple deep learning architectures were implemented and evaluated to identify the most effective model for freshwater fish disease classification. In particular, the following models were chosen because they have demonstrated their effectiveness in the task of image classification and because of their architectural diversity: DenseNet, MobileNetV2, InceptionV3, VGG16, and VGG19. These models have been selected through comparative studies that have been reported in earlier researches and they reveal that they are effective in deriving discriminative features of complex visual data. All architectures were implemented using state-of-the-art design principles and fine-tuned to suit the characteristics of the proposed dataset.

- ***DenseNet:*** DenseNet, or Dense Convolutional Network, introduces direct connections between all layers in a feed-forward manner. Unlike conventional convolutional neural networks (CNNs), where each layer is connected only to its immediate successor, DenseNet establishes connections from each layer to every other subsequent layer. As a result, a DenseNet with L layers contains $L(L + 1)/2$ direct connections, whereas a traditional CNN with L layers contains only L connections. In DenseNet, the output of each layer is concatenated with the outputs of all preceding layers and used as input for the subsequent layers. The layer transition can be expressed as:

$$X_l = H_l(X_l - 1) + X_l - 1 \qquad (1)$$

Where H_l represents a composite function consisting of convolutional, normalization, and activation operations.

This dense connectivity pattern enables efficient feature reuse, strengthens gradient flow, and significantly reduces the number of trainable parameters. Consequently, DenseNet alleviates the vanishing gradient problem and improves learning efficiency, particularly in deep architectures. DenseNet presents both the introduction of feature maps simultaneously with ResNet, which presents additive skip connections to maintain identity mappings, allowing a more detailed information flow and more ability to store learned features across layers. DenseNet (fig. 5) is especially appropriate in this regard because of its architectural design, which poses no problem with such tasks that are highly sensitive to feature discrimination, like fish disease classification

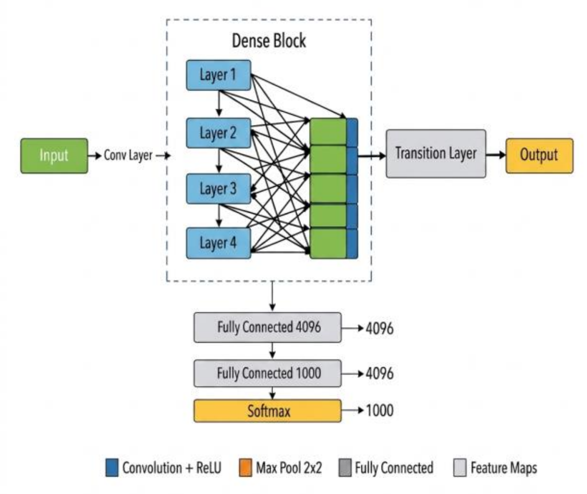


Fig. 5. A CNN model- DenseNet Architecture

- ***MobileNet V2:*** MobileNetV2 presents a new, better convolutional neural network architecture using inverted residual block and linear bottleneck layers, which are developed to ensure high performance but have low computation complexity. This architecture is particularly well-suited for mobile and embedded vision applications. MobileNetV2 builds upon the depth-wise separable convolutions introduced in MobileNetV1 by expanding feature representations within inverted residual blocks and projecting them back to a lower-dimensional space using linear bottlenecks. The core building block of MobileNetV2 consists of three main components: a pointwise convolution for feature expansion, a depth-wise convolution for spatial filtering, and a linear pointwise convolution for dimensionality reduction. This design enhances representational power while minimizing information loss and computational cost. As a result, MobileNetV2 achieves improved accuracy and faster inference compared to its predecessor, making it effective for a wide range of computer vision tasks, including image classification, object detection, and semantic segmentation. Fig. 6 represents the architecture of the model.

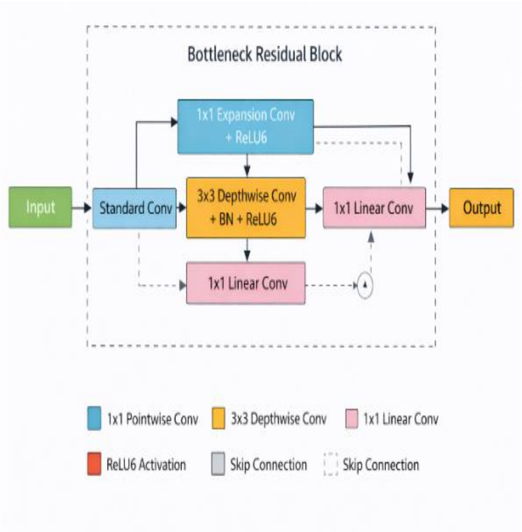


Fig. 6. A Fundamental architecture of a CNN model- MobileNet

- ***Inception V3:*** Inception-V3 is a deep convolutional neural network architecture belonging to the Inception family, designed to achieve high classification accuracy with improved computational efficiency. The architecture incorporates several advanced techniques, including label smoothing, factorized convolutions (such as decomposed 7×7 convolutions), and auxiliary classifiers. The auxiliary classifiers, combined with batch normalization in the side branches, help distribute label information to intermediate layers, improve gradient flow, and act as regularizers during training.

To effectively capture multi-scale spatial features, the Inception-V3 architecture is organized into three main stages: (1) five Inception-A modules, (2) four Inception-B modules, and (3) two Inception-C modules. The network employs grid-size reduction operations at specific points to downsample feature maps while preserving essential spatial and semantic information. During training, auxiliary classifiers are introduced at intermediate layers to enhance learning stability and mitigate the vanishing gradient problem.

Given an input image of size 299 × 299 × 3, Inception-V3 produces a final feature map of size 8 × 8 × 2048. This feature representation is flattened and passed through a fully connected layer with softmax activation for final classification into 1001 categories. Throughout the network, spatial dimensions are gradually reduced while the number of feature channels increases. By leveraging factorized convolutions, label smoothing, and batch normalization, Inception-V3 achieves greater computational efficiency and improved performance compared to earlier Inception architectures. The overall structure of the Inception-V3 model is illustrated in fig. 7.

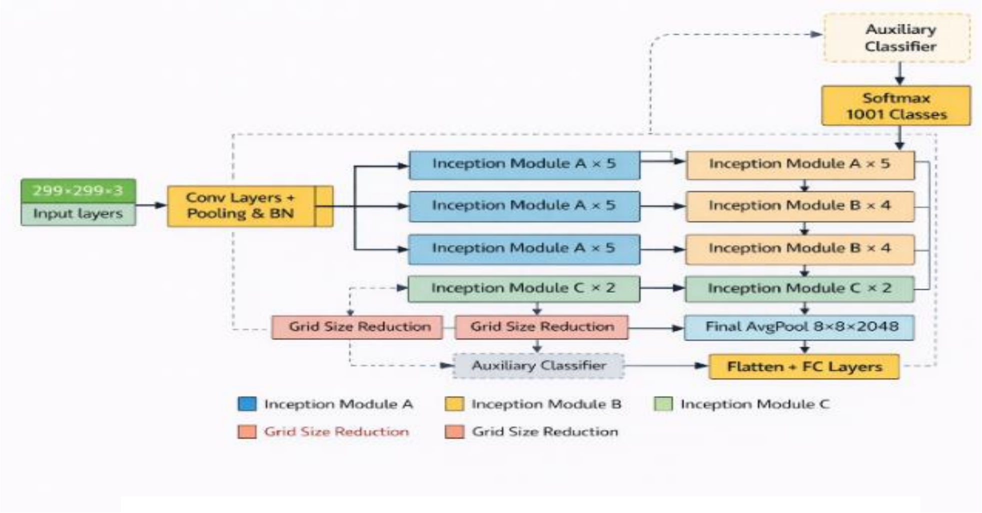


Fig. 7. An architecture of a CNN model Inception V3

- ***VGG16:*** VGG16 is a deep convolutional neural network architecture widely used for image recognition tasks due to its simple yet effective design. The model consists of 16 trainable layers, including 13 convolutional layers and 3 fully connected layers. Its architecture is characterized by the sequential stacking of small 3×3 convolutional filters, followed by Rectified Linear Unit (ReLU) activation functions to introduce non-linearity. Max-pooling layers with a kernel size of 2×2 are applied after every few convolutional layers to progressively reduce spatial dimensions while retaining the most significant features. After feature extraction, the resulting feature maps are flattened and passed through three fully connected layers. The final layer employs a softmax activation function to perform multi-class classification (fig.8).

For an input image of size 224 × 224 × 3, VGG16 gradually reduces the spatial resolution while increasing the depth of feature channels, ultimately producing a fully connected output with 1,000 class probabilities in its original configuration. Despite its strong feature representation and effectiveness in transfer learning applications, VGG16 is computationally expensive due to its large number of parameters, approximately 138 million.

- ***VGG19:*** The Visual Geometry Group (VGG) at the University of Oxford introduced the

VGG19 deep convolutional neural network architecture. The model comprises a total of 19 trainable layers, including 16 convolutional layers, 3 fully connected layers, and a final softmax output layer for classification. VGG19 follows a simple yet effective architectural design that relies exclusively on 3 × 3 convolutional filters, stacked with increasing depth, combined with ReLU activation functions and max-pooling layers for spatial down-sampling.

VGG19 has demonstrated strong performance in large-scale image classification tasks, such as ImageNet, due to its ability to learn highly discriminative feature representations. However, a major limitation of VGG-19 is, large number of parameters (approximately 143 million). Despite this drawback, VGG19 remains widely used for feature extraction, transfer learning, and style transfer applications, owing to its robust and hierarchical feature learning capability, as shown in the fig. 9.

- ***Vision Transformer (ViT):*** Vision Transformer (ViT) model is transformer learning based deep learning model. It' one kind of neural network architecture which focuses on the relationship within data to extract information more effectively. This transformer-based model uses attention mechanism to capture the similarities between inputs and process the entire sequential work at once. That's why it doesn't depend on convolution operation like others CNN model. ViT uses global similarities of input images by using self-attention mechanism instead of capturing local features like others convolutional network.

ViT reduces biasness but it needs large amounts of images for better performance. In our model, ViT is used to compare the result between others CNN models to observe the performance of low resource real world image of fresh fish. ViT architecture is demonstrated in fig. 10.

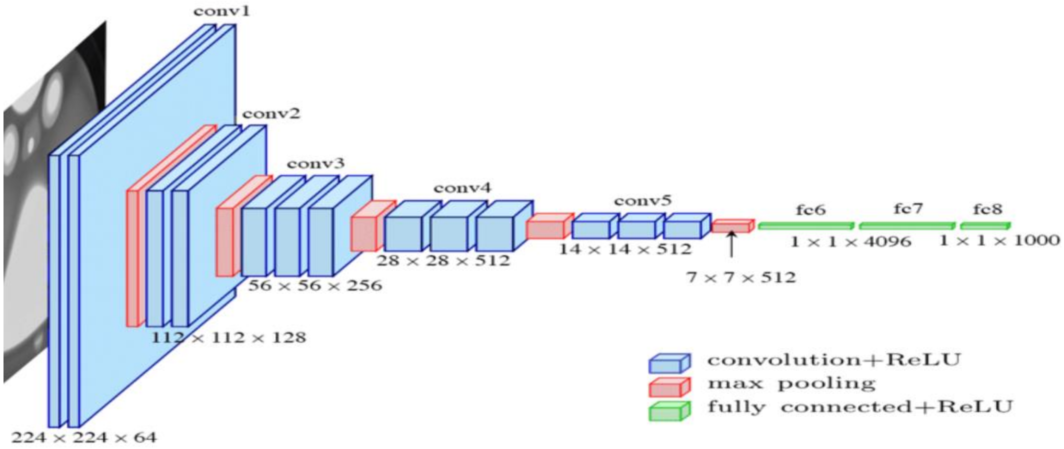


Fig. 8. An Architecture of A CNN model: VGG16

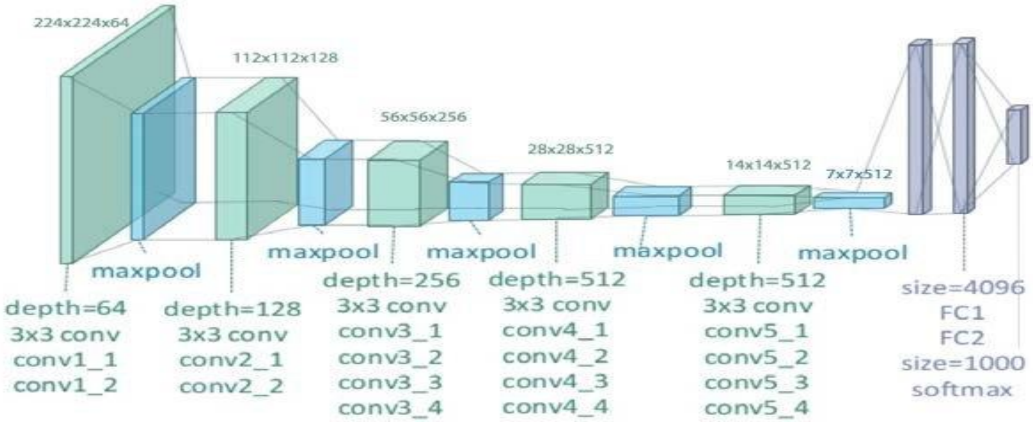


Fig.9. An Architecture of A CNN model: VGG-19

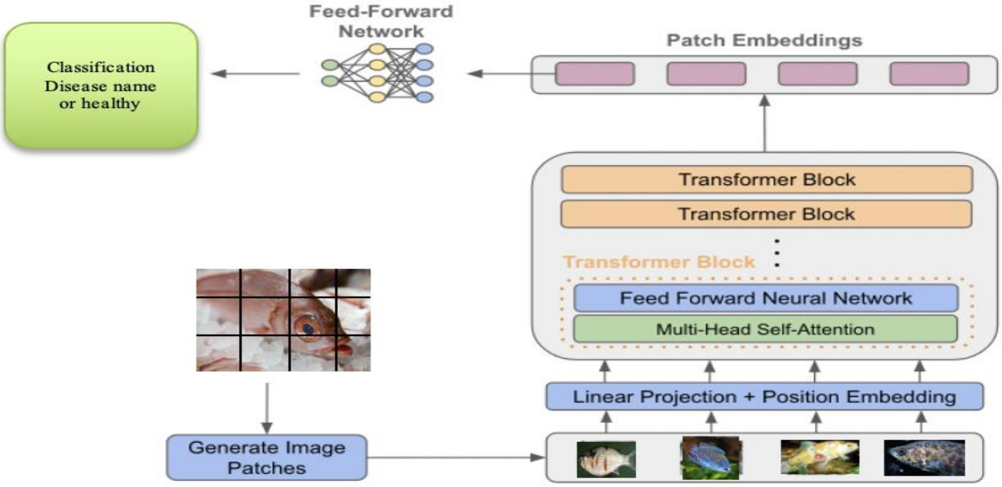


Fig.10. The Vision transformer (ViT) model architecture.

4. Result Discussion

This work achieved satisfactory classification accuracy using the raw freshwater real world fish image dataset. The effectiveness of the proposed approach is demonstrated across multiple evaluation dimensions. The performance metrics and analytical outcomes below support the validity of the obtained results, and each of them focuses on a particular aspect of the experimental analysis.

4.1. Performance Measurement Metrics:

The predictive performance of the individual convolutional neural network (CNN) models was measured using key performance metrics such as accuracy, precision, recall, and F1-score. Those measures allow carrying out a full quantitative evaluation of the models in terms of their capacity to correctly detect and categorize freshwater fish diseases. The results of the evaluation are significant information about the effectiveness and reliability of the offered approach that may be used in fisheries science and health management of aquaculture.

$$\text{Accuracy} = (TN+TP) / (TN+TP+FN+FP) \tag{2}$$

$$\text{Precision} = TP / (TP+FP) \tag{3}$$

$$\text{Recall} = TP / (TP+FN) \tag{4}$$

$$\text{F1 Score} = (2(\text{Precision}*\text{Recall})) / (\text{Precision} + \text{Recall}) \tag{5}$$

In this case, TP, TN, FP, and FN represent True Positive, True Negative, False Positive, and False Negative, respectively [32]. Besides that, the Area under the Curve (AUC) metric was also calculated to give a more detailed assessment of model performance. The next section is the analysis and discussion of the experimental results based on the implementation of the proposed models and the evaluation of their performance according to all the evaluation metrics.

4.2. Accuracy Analysis

The models of the convolutional neural network include DenseNet, InceptionV3, MobileNet, VGG16, and VGG19, are discussed and analyzed in this section. Model performance was determined by classification accuracy, which was followed by the evaluation of other metrics. The findings show that DenseNet and MobileNet had the best accuracy at 90%, which meant that they had a high ability to learn discriminative features and accurately detect diseases of freshwater fish using medical image data. Conversely, InceptionV3, VGG16, and VGG19 have accuracies of 89, 87, and 85 respectively and are slightly poorer than models that performed best. The additional investigation of the accuracy, recall, F1-score, and support of DenseNet and MobileNet proves the consistency and validity of their predictions, where such values remain stable and high in all classes. To train and test the model, the data was stratified into 80 percent training, 10 percent validation and 10 percent testing. Each of the models was trained in 10 epochs with a batch size of 75 images. The learning curves depict consistent and strong convergence behavior during the training process. Table 1 shows a summary of the overall performance of the five implemented CNN models, reports weighted average evaluation metrics to enable overall comparison.

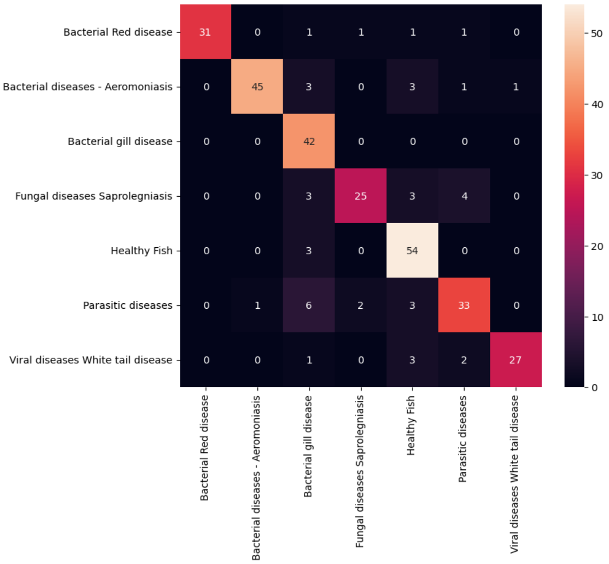
Table 1: CNN and ViT model Performance Matrix

Models	Precision	Recall	F1- Score	Accuracy
VGG 16	0.89	0.88	0.87	0.87
VGG 19	0.89	0.90	0.88	0.86
DenseNet	0.90	0.88	0.89	0.90
MobileNet	0.88	0.89	0.90	0.90
Inception V3	0.88	0.88	0.89	0.89
Vision Transformer (ViT)	0.78	0.79	0.78	0.79

The critical comparation of the accuracy of classification, as highlight in the table 1, is that the accuracy of the proposed models is up to 90%. Altogether, this comparative study has provided some useful information and can direct future studies on more specific, more reliable and scalable solutions to solve the chronic issues of freshwater fish disease monitoring. The high performance of DenseNet and MobileNet is explained by the fact that they have the right architectural features that are suitable for classifying freshwater fish diseases based on the image. DenseNet has the advantage of dense connectivity, and later layers reuse features of all previous layers. Such a design enhances the propagation of features, gradient flow and allows the model to learn more about minor disease-related visual details including lesions, discolouration and surface anomalies. In addition, DenseNet minimizes the redundancy of parameters, which is useful in avoiding overfitting in the case of moderately sized datasets. MobileNet uses depth-wise separable convolutions and inverted residual blocks with linear bottlenecks to enable the achievement of competitive accuracy. These elements enable the network to acquire discriminative features effectively and with minimal computational complexity. The lightweight architecture results in consistent convergence and better generalization as indicated by the well-matched training and validation loss curves. By contrast, VGG16 and VGG19 have much more parameters and do not have more sophisticated features reuse mechanisms, and thus, they are more susceptible to overfitting. InceptionV3 is a multi-scale feature capture model, but its complexity can hinder performance gains with limited training data.

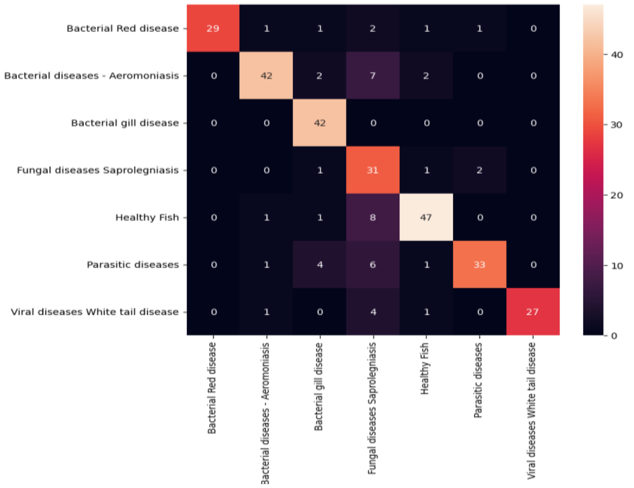
4.2.1. *Confusion Matrix*

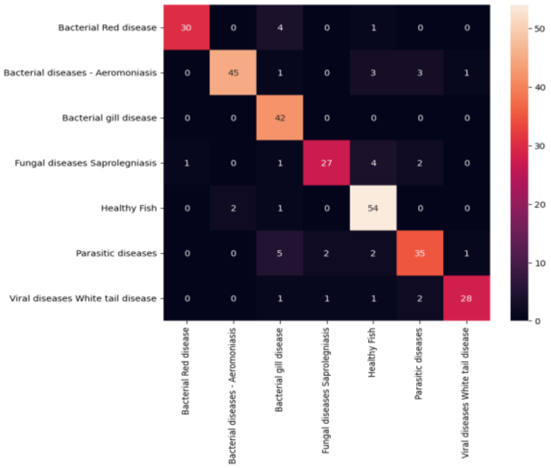
Fig. 5 demonstrates the validity of the classification results with the help of the confusion matrices. These matrices give a specific break down of the model predictions, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values of each class. Although the models have high true positive rates on the whole, certain misclassifications are found, especially in a few disease categories where there are relatively higher numbers of false positives and false negatives. Interestingly, the highest classification performance is achieved in Figures 5(d) and 5(e), which represent the DenseNet and MobileNet models. The real positive rates are high in such instances, and the false positive, and false negatives are relatively low; therefore, more accurate and confident predictions. Altogether, the confusion matrix analysis enables concluding that the models of DenseNet and MobileNet are more effective when compared to the rest of the evaluated ones regarding freshwater fish disease classification. To compare the performance of our CNN model, we present the confusion matrices of our CNN model in fig. 11.

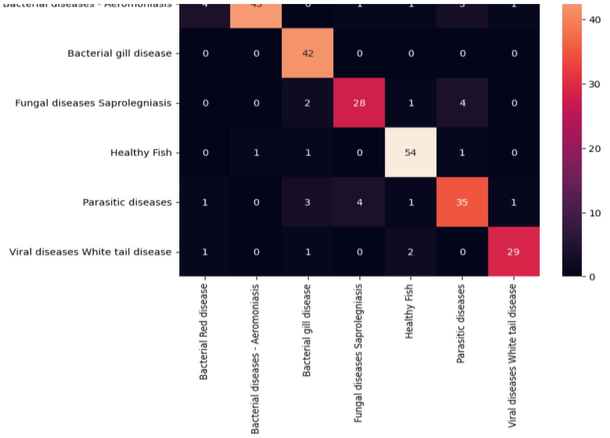


	Bacterial Red disease	Bacterial diseases - Aeromoniasis	Bacterial gill disease	Fungal diseases Saprolegniasis	Healthy Fish	Parasitic diseases	Viral diseases White tail disease
Bacterial Red disease	31	0	1	1	1	1	0
Bacterial diseases - Aeromoniasis	0	45	3	0	3	1	1
Bacterial gill disease	0	0	42	0	0	0	0
Fungal diseases Saprolegniasis	0	0	3	25	3	4	0
Healthy Fish	0	0	3	0	54	0	0
Parasitic diseases	0	1	6	2	3	33	0
Viral diseases White tail disease	0	0	1	0	3	2	27

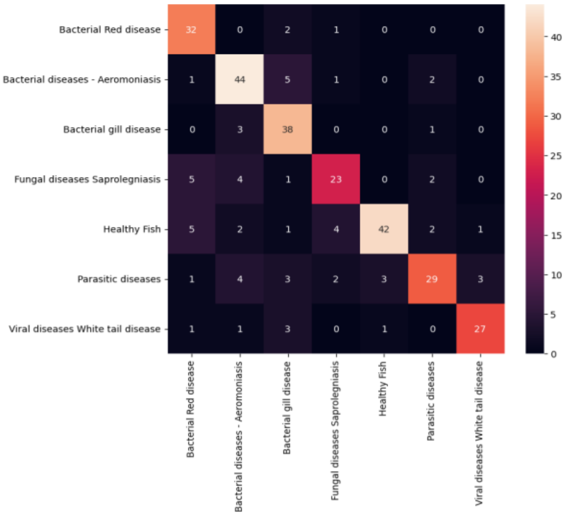
(a) Confusion Matrix of VGG16



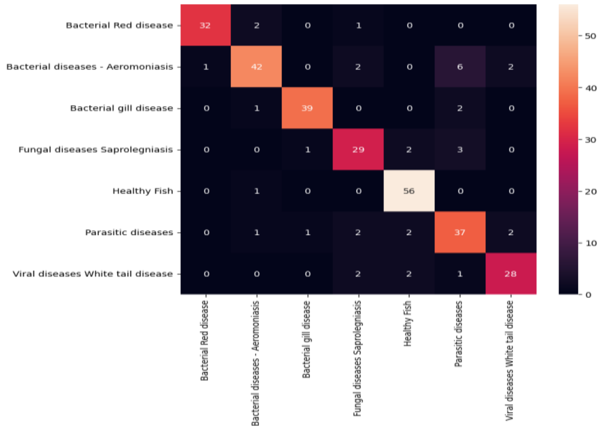




(c) Confusion Matrix of MobileNet



(f) Confusion Matrix of ViT



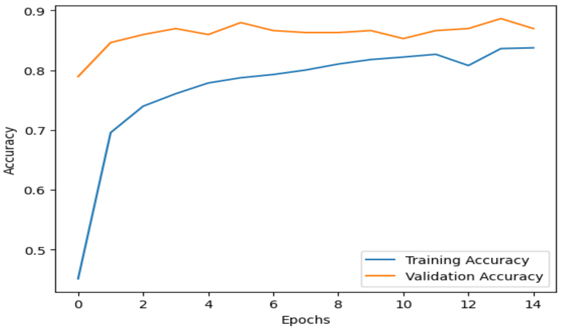
(d) Confusion Matrix of DenseNet

Fig. 11. Confusion matrices of our analyzed CNN models

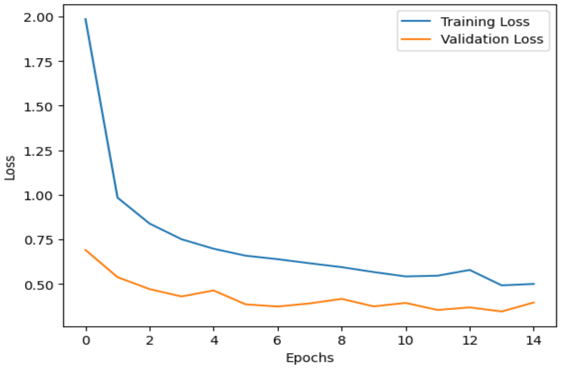
The analysis of the confusion matrices shows that both VGG16 and VGG19 models are characterized by rather high misclassification rates in terms of higher values of false positives (FP) and false negatives (FN). DenseNet and MobileNet, conversely, have higher true positive rates and are able to minimize misclassification among disease categories. On balance, the obtained results of the confusion matrix indicate that the chosen models can be used to classify freshwater fish disease with a high degree of reliability and satisfaction.

4.2.2. *Accuracy and Loss Curve*

The training and validation accuracy and loss curves are provided to give a better insight into the performance and reliability of the proposed models. The accuracy and loss curves of VGG16 and VGG19 are presented in Figures 12(a) and 12(b), respectively, and yield a final accuracy of 0.85 and 0.87, respectively. Even though these models show a stable convergence, their level of performance is relatively lower than the performance of the other architectures. Fig.12(c), fig.12(d), and fig.12(e) show how Inception V3, DenseNet, and MobileNet train their behavior, with their accuracy growing steadily without any apparent sign of overfitting. Moreover, the validation loss is also close to the training loss, which means good generalization ability. On the whole, the graphical analysis proves that InceptionV3, DenseNet, and MobileNet can work successfully with the raw dataset, and DenseNet and MobileNet demonstrate the best and the most stable learning behavior.

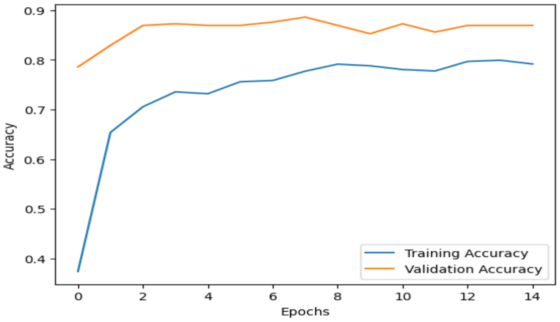


I. Accuracy Graph of VGG16

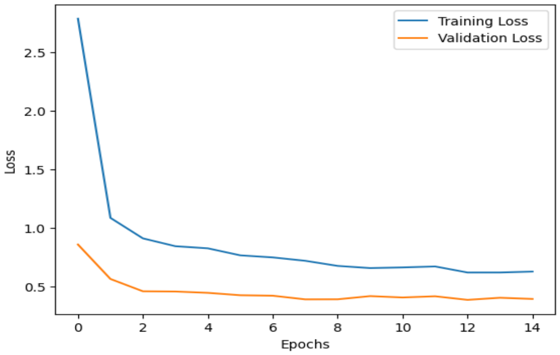


II. Loss Graph of VGG16

(a) VGG16 accuracy and Loss Graph

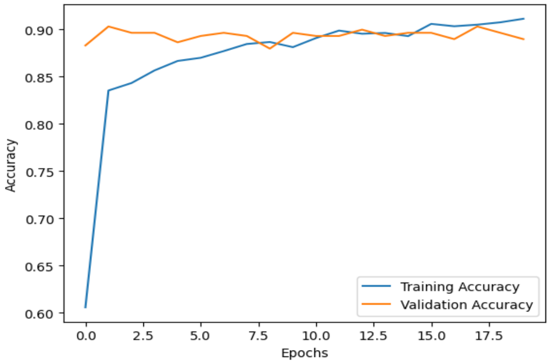


I. Accuracy Graph VGG19

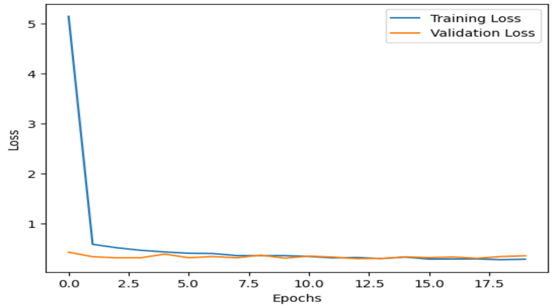


I. Loss Graph of VGG19

(b) VGG19 accuracy and loss graph

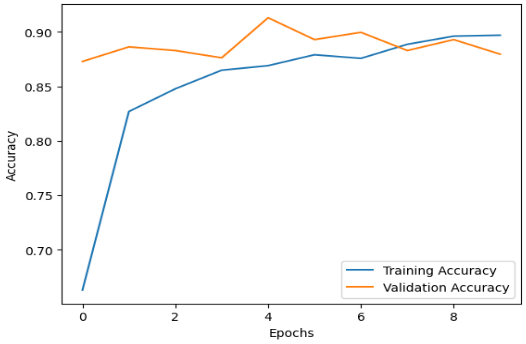


I. Accuracy Graph of DenseNet

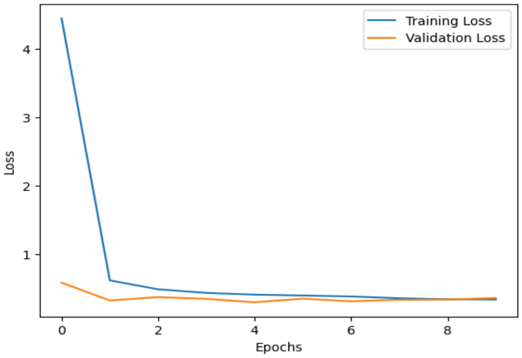


II. Loss Graph of DenseNet

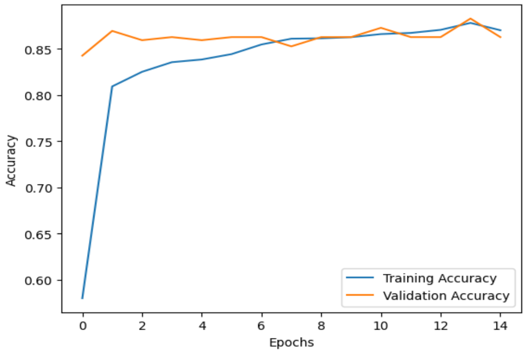
(c) DenseNet accuracy and loss graph



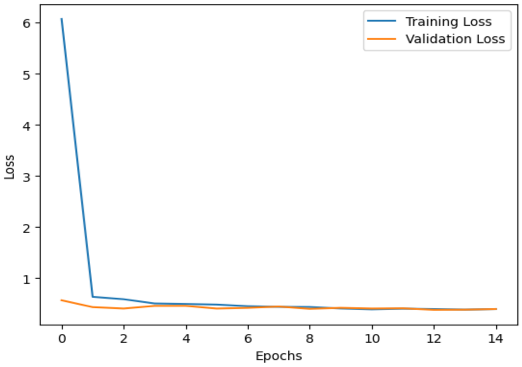
I. Accuracy Graph of MobileNet



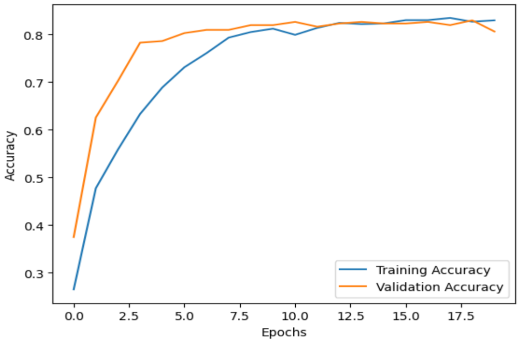
II. Loss Graph of MobileNet
(d) MobileNet accuracy and loss graph



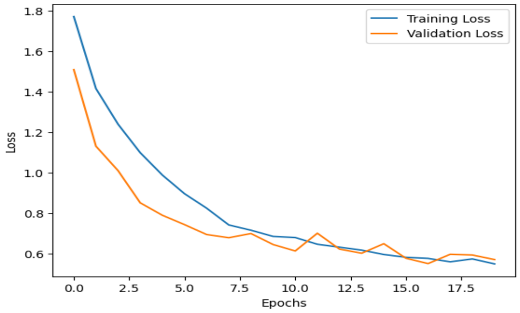
I. Accuracy Graph of Inception V3



II. Loss Graph of Inception V3
(e) Inception V3 accuracy and loss graph



I. Accuracy graph of ViT



II. Loss graph of ViT
(f) Vision Transformer (ViT) model accuracy and loss graph

Fig. 11. Accuracy and Loss Graph of our analyzed models.

4.2.3. *Comparative Analysis*

The comparative study between our model and other existing model is shown in table 2. The proposed work highlights on real environmental dataset rather than processed dataset which has been used in previous research.

Table 2: Comparative Evaluation of Fish Disease Detection Models Reported in the Literature

Study	Method	Dataset	Real Environmental image	Accuracy (%)
Ahmed et al. [1]	SVM	Salmon fish disease dataset	No	94.12

Study	Method	Dataset	Real Environmental image	Accuracy (%)
Sikder et al. [18]	M-SVM + GLCM	Freshwater fish diseases	No	88.00
Desai et al. [12]	Deep Learning CNN	Fish species dataset	No	100.00
Villon et al. [20]	CNN	Underwater fish images	No	94.90
Proposed Study	DenseNet	7-class disease of freshwater fish	Yes	90.00
Proposed Study	MobileNetV2	7-class disease of freshwater fish	Yes	90.00

Table 2 compares the proposed framework with several previously reported fish disease and fish classification studies. Although some studies achieved higher accuracies, many were conducted on smaller datasets, species classification tasks, or controlled image environments. In contrast, the proposed study focuses on seven classes of freshwater fish diseases using real-world images collected from publicly available repositories. DenseNet and MobileNetV2 achieved 90% classification accuracy while maintaining strong precision, recall, and F1-score values. These results demonstrate that the proposed framework provides competitive performance and is suitable for practical aquaculture disease monitoring applications.

5. Limitations and Future Work

In this research, the research practices were observed, such as the thorough literature review and the systematic data collection, to compare the CNN and ViT models for disease classification. The present paper is limited to six typical freshwater fish disease categories; it might be considered that, in the future, more disease classes with more sample diversity are to be included in the data set. The consideration of country-specific disease distribution and environmental variables, including climate and weather changes, can also be included to increase the range of applicability of the system and make more informed decisions related to aquaculture management. In this work, deep learning models were used with pretraining and modified by using preprocessing methods and the architecture of DenseNet as image classifiers, which demonstrated high-quality and reliable performance. However, more studies are needed to enhance the functionality and readability of models. As an illustration, the inclusion of disease localization methods to detect the diseased areas of fish images would offer more practical information. Attention-based and vision-based transfer learning are viable strategies towards the realization of this goal. Further on, the future work can be done on the usability and functionality of the developed web application to facilitate the wider adoption of the application by farmers, researchers, and aquaculture practitioners.

6. Conclusion

This study proposed a deep learning-based framework for detecting and classifying freshwater fish diseases from images. Six pre-trained CNN models: VGG16, VGG19, InceptionV3, DenseNet, and MobileNetV2 and ViT were evaluated using a dataset containing six freshwater fish diseases and healthy fish images. Among the tested models, DenseNet and MobileNetV2 achieved the best performance, reaching 90% classification accuracy along with high recall and F1-score values. A key contribution of this work is its focus on real-world image conditions. In practical aquaculture environments, fish images are often captured using mobile phones or cameras under varying lighting conditions and may contain noise or background interference. Unlike studies that rely on carefully selected images, this framework considers these challenges and applies appropriate preprocessing techniques to improve model robustness. As a result, the proposed approach is more suitable for real-world deployment and field-level disease monitoring. The findings demonstrate that deep learning can serve as an effective tool for automated fish disease diagnosis. Early disease detection can help reduce economic losses, improve fish health management, and support sustainable aquaculture practices. Future research may incorporate environmental factors, water quality parameters, and temporal data to develop a more comprehensive disease prediction system. Overall, this study highlights the potential of computer vision and deep learning technologies to provide practical and scalable solutions for freshwater fish health monitoring.

Acknowledgement

We are sincerely grateful to Daffodil International University - Department of CSE (Computer Science and Engineering), and Department of CIS (Computing and Information System) for all the assistance and advice they have provided in this project work. The Research Center of DIU approves the ethical nature of this research work.

References

[1] M. S. Ahmed, T. T. Aurpa, and M. A. K. 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, pp. 5170–5182, 2022.

[2] A. G. J. Tacon and M. Metian, "Fish matters: Importance of aquatic foods in human nutrition and global food supply," *Reviews in Fisheries Science*, vol. 21, no. 1, pp. 22–38, 2013.

[3] "Bangladesh achieves remarkable success in fish production," *Daily Industry News*. [Online]. Available: https://dailyindustry.news/bd-achieves-remarkable-success-in-fish-production/

[4] D. Kar, *Community-Based Fisheries Management*, Springer, New Delhi, India, 2021, pp. 121–505.

[5] C. R. Nichols and R. G. Williams, *Encyclopedia of Marine Science*, Infobase Publishing, New York, USA, 2009.

[6] M. Phillips, R. P. Subasinghe, N. Tran, L. Kassam, and C. Chan, *Aquaculture Big Numbers*, FAO Fisheries and Aquaculture Technical Paper, FAO, Rome, Italy, 2016.

[7] A. Mitra, R. Ghosh, A. Mallik, K. Mondal, S. Zaman, and K. Banerjee, "Brackish-water aquaculture: A new horizon in climate change matrix," in *Sensitivity of Mangrove Ecosystem to Changing Climate*, Springer, New Delhi, India, 2013, pp. 245–323.

[8] D. Li, X. Li, Q. Wang, and Y. Hao, "Advanced techniques for the intelligent diagnosis of fish diseases: A review," *Animals*, vol. 12, no. 21, p. 2938, 2022.

[9] FAO, *The State of World Fisheries and Aquaculture 2020*, Food and Agriculture Organization of the United Nations, Rome, Italy, pp. 1–244, 2020.

[10] K. D. Lafferty, C. D. Harvell, J. M. Conrad, C. S. Friedman, M. L. Kent, A. M. Kuris, and S. M. Saksida, "Infectious diseases affect marine fisheries and aquaculture economics," *Annual Review of Marine Science*, vol. 7, no. 1, pp. 471–496, 2015.

[11] B. T. Lunestad, L. Nesse, J. Lassen, B. Svihus, T. Nesbakken, K. Fossum, and S. Yazdankhah, "Salmonella in fish feed: Occurrence and implications for fish and human health in Norway," *Aquaculture*, vol. 265, no. 1–4, pp. 1–8, 2007.

[12] N. P. Desai, M. F. Baluch, A. Makrariya, and R. MusheerAziz, "Image processing model with deep learning approach for fish species classification," *Turkish Journal of Computer and Mathematics Education*, vol. 13, no. 1, pp. 85–99, 2022.

[13] D. Bau, J. Y. Zhu, H. Strobelt, A. Lapedriza, B. Zhou, and A. Torralba, "Understanding the role of individual units in a deep neural network," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30071–30078, 2020.

[14] Z. Yan, Y. Bi, B. Xue, and M. Zhang, "Automatically extracting features using genetic programming for low-quality fish image classification," in *Proc. IEEE Congress on Evolutionary Computation (CEC)*, Kraków, Poland, 2021, pp. 2015–2022.

[15] M. A. Iqbal, Z. Wang, Z. A. Ali, and S. Riaz, "Automatic fish species classification using deep convolutional neural networks," *Wireless Personal Communications*, vol. 116, pp. 1043–1053, 2021.

[16] A. Jalal, A. Salman, A. Mian, M. Shortis, and F. Shafait, "Fish detection and species classification in underwater environments using deep learning with temporal information," *Ecological Informatics*, vol. 57, p. 101088, 2020.

[17] S. Malik, T. Kumar, and A. K. Sahoo, "Image processing techniques for identification of fish disease," in *Proc. IEEE 2nd Int. Conf. on Signal and Image Processing (ICSIP)*, Singapore, 2017, pp. 55–59.

[18] J. Sikder, K. I. Sarek, and U. K. Das, "Fish disease detection system: A case study of freshwater fishes of Bangladesh," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 867–871, 2021.

[19] A. A. Nayan, A. N. Mozumder, J. Saha, K. R. Mahmud, A. K. A. Azad, and M. G. Kibria, "A machine learning approach for early detection of fish diseases by analyzing water quality," *arXiv preprint*, arXiv:2102.09390, 2021.

[20] S. Villon, D. Mouillot, M. Chaumont, E. S. Darling, G. Subsol, T. Claverie, and S. Villéger, "A deep learning method for accurate and fast identification of coral reef fishes in underwater images," *Ecological Informatics*, vol. 48, pp. 238–244, 2018.

[21] A. Tharwat, A. A. Hemedan, A. E. Hassanien, and T. Gabel, "A biometric-based model for fish species classification," *Fisheries Research*, vol. 204, pp. 324–336, 2018.

[22] Z. Wang, H. Liu, G. Zhang, X. Yang, L. Wen, and W. Zhao, "Diseased fish detection in the underwater environment using an improved YOLOv5 network for intensive aquaculture," *Fishes*, vol. 8, no. 3, p. 169, 2023.

[23] G. Yu, Y. Zhou, J. Li, and X. Sun, "Detection and identification of fish skin health status using deep learning," *Oceans*, vol. 8, no. 4, pp. 186–201, 2023.

[24] A. Al Muksit, F. Hasan, M. F. H. Bhuiyan, M. R. Haque, A. R. Anwary, and S. Shatabda, "YOLO-Fish: A robust fish detection model for underwater environments," *Ecological Informatics*, vol. 73, p. 101847, 2022.

[25] Z. Wang, Y. Liu, H. Zhang, and X. Chen, "Diseased fish detection in underwater environments using deep learning," *Oceans*, vol. 8, no. 3, pp. 169–184, 2023.

[26] B. Austin, D. A. Austin, and C. B. Munn, *Bacterial Fish Pathogens: Disease of Farmed and Wild Fish*, Springer, Dordrecht, The Netherlands, 2007.

[27] R. C. Cipriano, *Aeromonas hydrophila and Motile Aeromonad Septicemias of Fish*, U.S. Fish and Wildlife Service, Washington, DC, USA, 1984.

[28] E. J. Wagner, R. E. Arndt, M. D. Routledge, and Q. Bradwisch, "The effects of formalin, hydrogen peroxide, and sodium chloride on bacterial gill disease and costia," *North American Journal of Aquaculture*, vol. 64, no. 4, pp. 248–256, 2002.

[29] D. W. Bruno and B. P. Wood, "Saprolegnia and other Oomycetes," in *Fish Diseases and Disorders*, vol. 3, P. T. K. Woo and D. W. Bruno, Eds., CABI Publishing, 1999, pp. 599–659.

[30] J. C. Eiras, H. Segner, T. Wahli, and B. G. Kapoor, *Fish Diseases*, vols. 1–3, Science Publishers, Enfield, USA, 2008.

[31] K. Wolf, *Fish Viruses and Fish Viral Diseases*, Cornell University Press, Ithaca, NY, USA, 1988.

[32] M. A. Islam, M. M. Rahman, A. A. Shohan, R. A. Pulok, S. Akter, and M. T. Ahmed, "Deep neural network-based approach to identify the paddy leaf disease using ResNet50-V2," in *Proc. Int. Conf. on Evolutionary Algorithms and Soft Computing*

Tapasy Rabeya is a Senior Lecturer at the Department of Computer Science and Engineering, Daffodil International University, Bangladesh, where she has been a faculty member since 2019. In 2018, she served as a Research Associate of the department. Rabeya graduated with a first-class honors B.Sc. degree in Computer Science and Engineering, Daffodil International University, Bangladesh, in 2017, and an M.Sc. in Computer Science and Engineering, Daffodil International

University, Bangladesh in 2021.Her research interests are primarily in the area of data mining, machine learning, NLP and computer vision, where she is the author/co-author of over 18 research publications.

Sonia Nasrin is a Lecturer at the Department of Computing and Information System, Daffodil International University, Bangladesh. She completed her B.Sc. degree in Computer Science and Engineering, Daffodil International University, Bangladesh and an M.Sc. in Computer Science and Engineering, Daffodil International University, Bangladesh. She is enthusiastic about artificial intelligence, machine learning, image processing and natural language processing.

Lamia Rukhsara received Bachelor of Science in Computer Science and Engineering from Military Institute of Science and Technology, Bangladesh and M.Sc. from Jahangirnagar University, Bangladesh. Currently, she is working as a senior lecturer at Department of Computer Science & Engineering, Daffodil International University, Bangladesh. Her research interests are in machine learning, data mining, cloud computing, health informatics and agricultural automation.

Israt Jahan is a dedicated and passionate University Lecturer (Senior Scale) of the Department of Computer Science and Engineering at Daffodil International University with a diverse academic background and a wealth of experience in teaching and education. Her interested research fields are Machine Learning, Deep Learning, Digital Image Processing, Computer Vision etc. She continues to make meaningful contributions to the academic community, leaving an indelible mark as a dedicated educator, researcher, and advocate for the transformative power of education.

Umme Habiba is currently serving as a Lecturer in the Department of Computer Science and Engineering (CSE) at Daffodil International University. She completed her Bachelor of Science (B.Sc.) in Computer Science and Engineering from the same institution and is currently, pursuing a Master of Science (M.Sc.) degree with a major in Data Science. Before, joining academia, she worked as a Web Developer at an IT company, where she gained practical experience in software development and web-based applications. Her research interests are in Machine Learning, Deep Learning, Image Processing, Health Informatics, and Agricultural Automation. She is particularly interested in developing intelligent and data-driven solutions to address real-world challenges and contribute to advancements in technology and society.