A COMPARATIVE ANALYSIS OF DEEP LEARNING APPROACHES FOR FISH DISEASE IDENTIFICATION

BY

MOBASSERA ASMA SADIA
ID: 221-25-146

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering

Supervised By
Dr. S. M. Aminul Haque
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By
Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Thesis titled **"A Comparative Analysis of Deep Learning Approaches for Fish Disease Identification"**, submitted by Mobassera Asma Sadia, ID No: 221-25-146 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17-01-2023.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori, PhD — **Chairman**
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Ms. Naznin Sultana — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mr. Md. Sadekur Rahman — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin, PhD — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Dr. S. M. Aminul Haque, Associate Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. S. M. Aminul Haque
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

Sadia

Mobassera Asma Sadia
ID: 221-25-146
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

At first, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project successfully.

We really grateful and wish our profound our indebtedness to **Dr. S. M. Aminul Haque, Associate Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Deep Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft, and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Professor Dr. Touhid Bhuiyan, Professor & Head,** Department of CSE, for his motivation and appreciation. We are also very thankful to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Bangladesh's fisheries and aquaculture industries play a significant role in the nation's food production, ensuring the food supply's nutritional security, growing agricultural exports, and employing 17 million people across various occupations. Farmers who farm fish face a lot of economic losses every year because of various diseases that can happen to fish. There are three common diseases of fish. They are known as black spots, red spots, and white spots. A parasite causes black spot disease. Red spot disease is also known as Epizootic ulcerative syndrome (EUS). And it is caused by a fungus. The white spot syndrome virus causes white spot disease. If a fish farmer can detect these diseases early and apply appropriate treatment, it may protect much infected fish and prevent economic loss. The manual approach of human visualization is a laborious effort for detecting and monitoring fish disease. As a result, any viable strategy that is quick, accurate, and highly automated encourages interest in this problem. Due to a lack of information and a high level of competence, there hasn't been a single piece of useful research on the fish disease. Our system provides solutions to this problem. Fish disease identification using deep learning from images is an arduous task. This study proposes a multi-classification deep learning model for identifying fish diseases (black spots, white spots, red spots, and healthy) from images. To classify and identify these diseases, we will apply five different pre-trained models (DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16), and we have also compared their accuracy. According to the experiment data, the MobileNetV2 model performed better than the other proposed models. In comparison to other models, the model provided good detection accuracy.

Keywords— Fish disease identification, Classification, TensorFlow, Dataset, Deep learning.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Introduction

Fisheries represent the potential for future expansion in Bangladesh's agricultural sector since they play a crucial part in social and economic evolution. Additionally, it plays a significant role in the world's food security. Fishing is one of Bangladesh's most active and productive sectors. The demand for fish keeps rising as the world's population rises and the benefits of fish as a source of animal protein are recognized. The fisheries industry employs over 17 million people, and fish alone provides about 60% of all animal protein. Almost 11 million people indirectly rely on fisheries-related activities for their livelihood. The export of fish and fish-related items generated 1.5% of the country's foreign exchange earnings, while the fishing sector contributed 3.57% to the national GDP and 25.30% to the agricultural GDP [1].

Scientists estimate that there are almost 22,000 kinds of fish in the globe. The demand for fish is growing as the world's population rises and as more people are becoming aware of the benefits of fish as an animal protein source. Fish consumption has notably increased recently in both developed and developing countries. In order to meet the rising demand for fish, the aquaculture industry is becoming more and more well-known as a more environmentally friendly method of ensuring a steady supply. The aquaculture industry must be both economically and environmentally efficient.

Fish diseases are one of the issues the fishing industry faces. Since fish carry infections and diseases like all living things do, they are susceptible to various ailments. Seasonal variations in water temperature and water contamination play a significant role in the propagation of fish diseases. Because it spreads quickly through the water, the fish disease is a severe problem for fishermen. The skin and gills of diseased fish must undergo an invasive and time-consuming histological study to make a diagnosis. Traditional methods for diagnosing fish infections involve an experienced fisherman or a fishery department specialist. On the other hand, a person's talent, experience, and understanding ultimately determine how accurate a final diagnostic of this kind is. This suggests that the final diagnosis can be inaccurate.

Artificial intelligence, which has validated its effectiveness and excellent performance in automatic image categorization problems using various machine learning algorithms, is currently being used to automate the detection of numerous diseases. In addition, machine learning refers to models that can learn from and decide based on a significant number of incoming data samples. Deep learning is a machine learning technique that primarily automatically extracts and classifies image features. It has demonstrated outstanding success in numerous applications. Without human intervention, deep learning effectively creates models that lead to more accurate predictions and classifications of various fish diseases using images of fish. The essential advantage of employing deep learning is that as the network becomes more complex, deep learning approaches learn by building an abstracter representation of the data. Because of this, the model automatically extracts information and produces results with more accuracy. In contrast to conventional machine learning algorithms, deep learning methods specify components through a collection of nonlinear functions used combinatorial to increase the model's accuracy.

This project introduces a multi-classification model based on deep learning techniques for detecting fish disease from images. Each of the architectures is thoroughly described in the study, and from the outcomes, it is determined which architecture is optimal and can reach the highest detection accuracy. Moreover, we compare deep-learning architectures using datasets for four classes of fish diseases: Black Spot, White Spot, Red Spot, and Healthy.

1.2 Motivation

Fish disease detection technologies have been analyzed over the past decades. Deep learning techniques can use large datasets of fish disease and learn costly and detailed representations of fish disease, allowing sophisticated models to perform well at first and then their ability to detect fish diseases afterwards. Even though deep learning methodology has advanced quickly, detecting infected fish still takes time. We want to create a model that can identify fish infections fast. We used some pre-trained models from Keras for Transfer Learning to construct a deep-learning network to address the problems of fish disease diagnosis. We deploy an automatic detection strategy with deep learning models. To achieve higher prediction accuracy, we have used some pre-trained models of DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16.

1.3 Objective

These diseases are spreads quickly. To handle the problem, we are presently working. First, recognize the fish disease known as Black Spot. Next, identify any fish that have Red Spot disease. Third, recognize the fish disease known as White Spot. Identify healthy fish, fourth. Finally, stop fish diseases from spreading. And raise the sector's productivity in the fisheries.

1.4 Expected Outcome

- Using classification to identify fish disease.
- Utilizing the Keras API, we will train the deep learning model using five different pre-trained models (DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16).
- Lastly, determine the accuracy of the suggested algorithms and compare them.
- Finally, with the help of this technology, we will be able to identify fish diseases quickly.

1.5 Report Layout

This report is organized as follows: The first chapter's introduction goes into more detail and provides a description of the project title. This section consists of the project's introduction in 1.1, the project's motivation in 1.2, and the project's objective in 1.3., the Expected outcome of the project in 1.4, and the report layout in 1.5. The second chapter, Background, discusses previous similar studies. Terminologies in 2.1, previous research activity or related works in 2.2, research analysis or a summary in 2.3, and challenges in 2.4 constitute this part. The research methodology is justified in chapter three. This chapter has discussed the project approaches in 3.1; deep learning proposed models in 3.2, a description of fish diseases in 3.3, a recommended processing model in 3.4 and model evaluation in 3.5 at the end of the section. The essential section of chapter four is the one given to materials, experiment results, and discussion. Here, the data set was utilized to build the project in 4.1, pre-processed in 4.2, experiment results and analysis shown in 4.3, results of several models compared in 4.4, and discussion presented in 4.5. In sections 5.1, 5.2, 5.3, and 5.4 of chapter 5, it is discussed how the environment, ethical considerations, and sustainability plan are impacted. The last section of chapter 6 provides an overview, a conclusion, and an analysis for future work in sections 6.1, 6.2, and 6.3, followed by the required references.

CHAPTER 2

Background

2.1 Terminologies

Fish disease detection is the process of identifying diseases from images of fish. In our case, we have used infected and healthy fish to detect fish disease. Several algorithms are used in some studies on it, although they don't always produce reliable results. We found that no research had been done to identify three diseases and healthy fish. A multi-classification deep learning model is created in this study to identify fish disease from images. To the best of our knowledge, this work is the first to identify fish diseases. Here, we deploy an automatic detection strategy with deep learning models to determine fish diseases. In this way, we will achieve higher prediction accuracy.

2.2 Related Works

The potential for future development in Bangladesh's agricultural sector should go to the fishing industry because it is important to social and economic development. Additionally, it plays a significant role in the world's food security. Fishing is one of Bangladesh's most active and productive sectors. Fish is increasingly in demand as the world's population rises and the benefits of fish as a source of animal protein become more widely recognized. The fisheries industry employs over 17 million people, and fish alone provides about 60% of all animal protein. Almost 11 million people indirectly rely on fisheries-related activities for their livelihood. The export of fish and fish-related items generated 1.5% of the country's foreign exchange earnings, while the fishing sector contributed 3.57% to the national GDP and 25.30% to the agricultural GDP.

Md. Mostafa Shamsuzzaman et al. [1] describe the performance of Bangladesh's fisheries. They discover that Bangladesh's fisheries sector has several difficulties, including the degradation of fisheries resources and overfishing. They propose recommendations for enhancing Bangladesh's fisheries' contribution to the national economy.

In 2021, a research paper was published about Fish Disease Detection in Aquaculture. In aquaculture, fish diseases are a serious threat to the security of the nutriment supply. By combining image processing and machine learning techniques, they are able to detect salmon fish disease in aquaculture and identify the pathogen-infected fish. They apply the Support Vector Machine (SVM) method to their work and receive accuracy results of 91.42 and 94.12 percent, with and without augmentation, respectively [2].

Juel Sikder et al. [3] suggested an approach based on image processing that automatically detects several fish diseases in freshwater, particularly in Bangladesh's Rangamati Kaptai Lake and Sunamganj Hoar region. The suggested system showed how K-means clustering and C-means fuzzy logic compare. The proposed methodology provided the highest accuracy rate compared to other current methods, with a K-means accuracy of 96.48% and a C-means accuracy of 97.90%.

Shaveta Malik et al. [4] proposed a study that the presence of epizootic ulcerative syndrome (EUS), a debilitating condition that affects freshwater and brackish water fish, has been rising in South-East Asia, Australia, Africa, Japan, and South-East Asia. According to the study, performed multiple edge recognition techniques in MATLAB software to actual images of fish infected with EUS.

Noraini Hasan et al. [5] demonstrates fish disease detection in the fish sector. They only used Convolutional Neural Network (CNN) in their analysis. They used only 90 samples for training and testing purposes; this is not enough to determine correct detection performance.

Jeong-Seon Park et al. [6] aim to achieve rapid disease diagnosis using image processing techniques from microscopic images of diseased fish parasitized. The limitation is that they only used microscopic images, which needed to give better accuracy, and it's a time-consuming process to fish direction.

Brian Austin et al. [8] proposed methods for fish-affecting bacterial and fungal diseases and poor water quality, together with poor water quality, are the most significant factors in fish fatalities. They did not use Deep Learning or Machine Learning in their analysis; they only used culture-independent approaches. Its limitations include a lack of sensitivity, a propensity for slowness, and a success rate depending on the media's creation and the incubation conditions used.

Md. Jueal Mia et al. [11] conducted a detailed analysis of expert systems that can proceed with an image obtained using a smartphone and identify the disease. The limitation is that they only used two sets of features, which was not giving good accuracy.

Teh Hong Khai et al. [20] proposed methods for counting shrimp underwater; this work effectively created a detection and recognition model based on a deep learning-based Mask R-CNN model. For the purpose of training the model, they employed a robotic eye camera to take pictures of shrimp on a shrimp farm. The accuracy rate of the suggested methodology is up to 97.48%.

Daoliang Li et al. [22] describes how image-processing technology is used in aquaculture, particularly for the diagnosis of fish diseases using images. The most significant technical methods for automatic image-based diagnosis and expert systems are outlined, and the future development of each diagnostic process is analyzed.

Abinaya N.S. et al. [23] proposed a method to combine multisegmented image-based DLNs for fish classification in aquaculture industries. The fish body, head, and scales were the three segments they used to train the DLNs using the transfer learning methodology. These DLNs are fused, and a naive Bayesian layer is added to improve classification accuracy. Their test results showed that the classification accuracy was 98.64% for the "Fish-Pak" image dataset and 98.94% for the BYU fish dataset.

We found that no research had been done to identify three diseases and healthy fish. All of the techniques mentioned depend entirely on machine learning and image processing.

2.3 Research Summary

In this study, we showed the way to fish disease identification. We have taken data from Kaggle and google sites. We develop a model that can identify fish disease promptly. We apply the Keras API framework to construct a deep-learning network that solves this problem. We accomplish this via a deep learning-based automatic detection strategy. We run different deep-learning techniques for detecting fish disease from images. Each of the architectures is thoroughly described in the study, and from the outcomes, it is determined which architecture is optimal and can reach the highest detection accuracy. Moreover, we compare deep-learning architectures using datasets for four classes of fish diseases. diseases.

2.4 Challenges

- Not having enough data to perform.
- Quantity of training data is small.
- Inconsistent training data.
- Data of poor quality.

CHAPTER 3

Research Methodology

3.1 Deep Learning Architecture

Deep Learning is a subset of Artificial Intelligence, which on the other hand is a subset of Machine Learning. Artificial intelligence (AI) is a broad phrase that refers to techniques that enable computers to replicate human behavior. Machine learning, which consists of a collection of algorithms trained on data, allows all this. On the other hand, Deep Learning is only a division of machine learning inspired by how the human brain is structured. Deep learning algorithms continually evaluate the data using a predetermined logical structure to reach conclusions comparable to those run by humans. Deep learning achieves this via neural networks, which are multi-layered structures of algorithms. Instead of being a single strategy, deep learning is a group of topologies and algorithms that can solve various problems.

Deep learning focuses on rewarding neural networks for success by training them with data; the more data available, the better it is for creating these deep learning structures. Deep learning uses a large and wide range of architectures and techniques. An input layer includes raw data; hidden layers contain essential information; hidden layers mix and process input data; and an output layer is the three main layers that make up neural networks (it delivers the outcome: result).

3.1.1 Convolutional Neural Network

Many applications of NLP, video analysis, image processing, and image recognition use convolutional neural networks. For tasks based on image categorization for fish disease prediction, complex artificial intelligence models based on a deep transfer learning technique are especially useful. Convolutional Neural Networks are able to process an input image, sort its various objects and elements into categories based on their significance, and identify one object from another. Convolutional refers to a mathematical operation where several functions are combined. Convolutional neural networks are made up of input, output, and several hidden layers. The majority of the hidden layers in convolutional neural networks are composed of convolutional layers.

Convolutional Neural Networks operate as follows: The network first receives the input. Each input image will pass through several convolution layers with various filters. The control layer guides the signal's transition between layers. The output then needs to be compressed before being sent into the fully connected layer, which connects every neuron in the network's layers to every neuron in the layer above it. The output can then be categorized as a result.

3.2 Deep Learning Proposed Models

In this study, we deal with the fish disease dataset. The proposed five classification models are built using a variety of supervised deep-learning approaches. We want to analyze how well they did at identifying the fish diseases that were under consideration and determine which one was the best. Each of these models combines CNN and RNN in a certain way. Five pre-trained models, DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 are used with the gated recurrent unit (GRU), standard CNN, and bidirectional gated recurrent unit (Bi-GRU) as types of RNN. The details of each of the five developed models are given in the following subsections.

3.2.1 DenseNet121

The complete form of DenseNet is Densely Connected Convolutional Networks. We choose this model because DenseNets have several effective benefits: they solve the vanishing gradient issue, promote feature reuse, improve feature propagation and drastically reduce the number of parameters. We used the pre-trained model DenseNet-121 in our study. To improve our model and avoid overfitting issues due to the limited image dataset, we blocked a number of the convolution layers. 120 Convolutions and 4 AvgPool are present in DenseNet-121. Since all layers are contained within the same dense block and transition layers disperse their weights over several inputs, hidden layers might utilize features that were previously collected. This model will receive as input images of fish disease with a dimension of (224, 224,3). To improve outcomes for a certain dataset, it tries to minimize the provided loss function. In Figure 3.1, the DenseNet-121 architecture is shown.

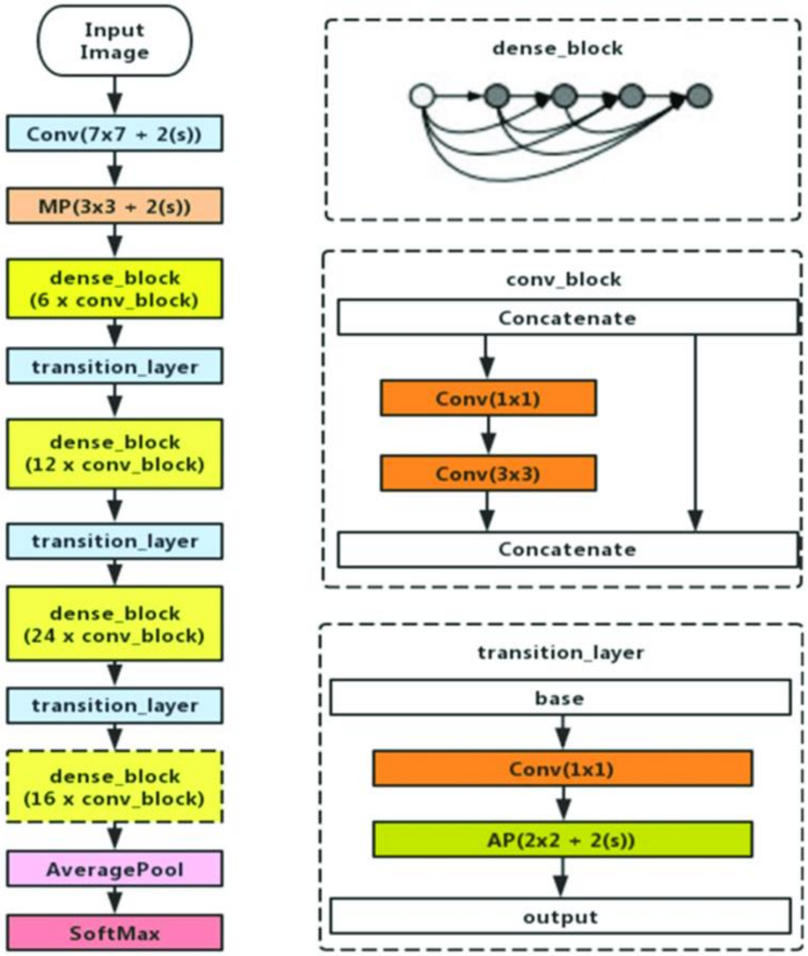

Fig 3.1: DenseNet-121 Architecture [26]

3.2.2 MobileNetV2

MobileNetV2 is used as a feature extraction model, and this is another deep model. Multiple use cases can be accommodated by MobileNetV2, which is a general architecture. The primary distinction between MobileNetV2 and the original MobileNet is the usage of inverted residual blocks with bottlenecking characteristics in MobileNetV2. When compared to the original MobileNet, there are much less parameters. Up to 32 × 32 pixels can be inputted with MobileNets, while larger image sizes function better. In Figure 3.2, the MobileNetV2 architecture is shown.

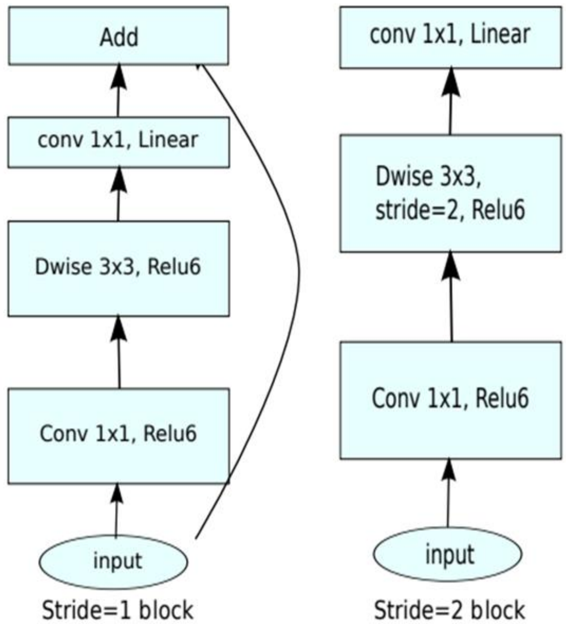

Fig 3.2: MobileNetV2 Architecture [27]

3.2.3 ResNet101V2

ResNet101V2 is a collection of ResNet building components that has 101 layers. Because it is a pre-trained model, ResNet101V2 can achieve acceptable accuracy more quickly than a typical CNN according to its initial weights. The input size for the system image is 224×224. In Figure 3.3, the ResNet101V2 architecture is shown.

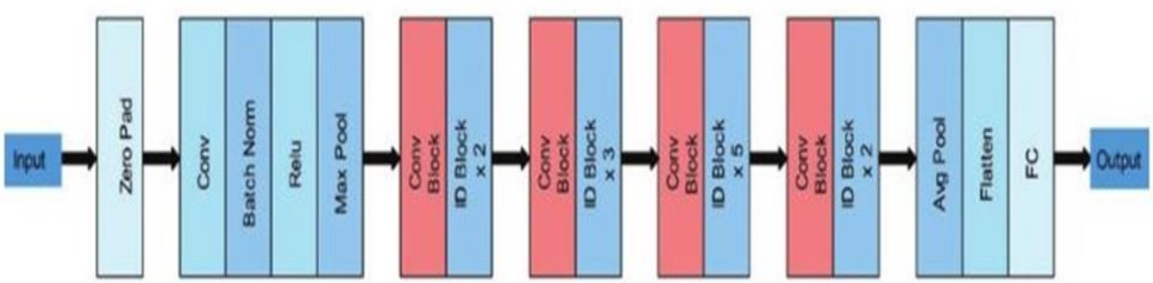

Fig 3.3: ResNet101V2 Architecture [28]

3.2.4 ResNet152V2

As a feature extraction model, the ResNet152V2 model is employed. The ResNet152V2 model serves as the framework for the model architecture, which also includes layers for flattening and reshaping, dense layers with 128 neurons, dropout layers, and dense layers with Softmax activation functions for categorizing images. The input size for the system image is 224×224. In Figure 3.4, the ResNet152V2 architecture is shown.

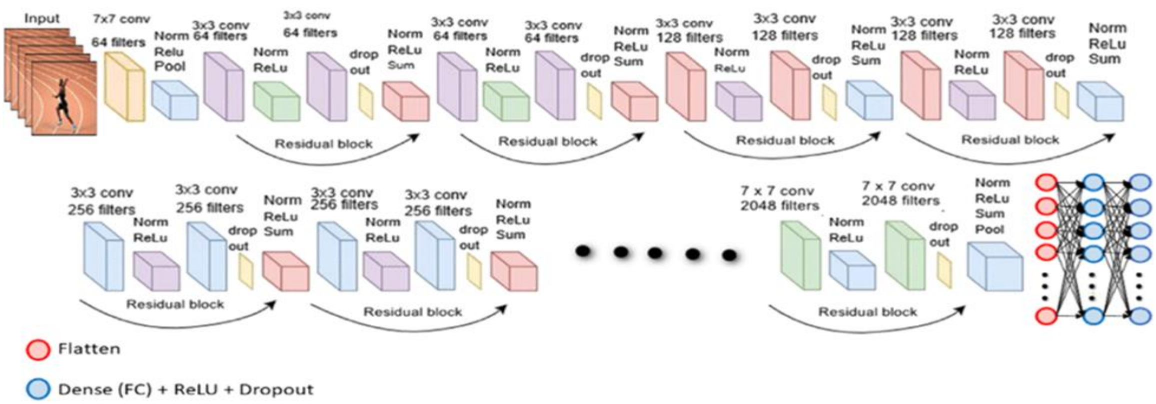

Fig 3.4: ResNet152V2 Architecture [29]

3.2.5 VGG16

We used the VGG16 pre-trained deep model in our experimentation. Due to small number of images in the dataset and to avoid overfitting problems, we blocked several convolutional layers to create a stronger model. This model's default input size is 224x224. A softmax output layer and two completely connected layers with ReLU activation functions are also included. A large network with about 138 million hyperparameters is the VGG-16 model. In Figure 3.5, the VGG16 architecture is shown.

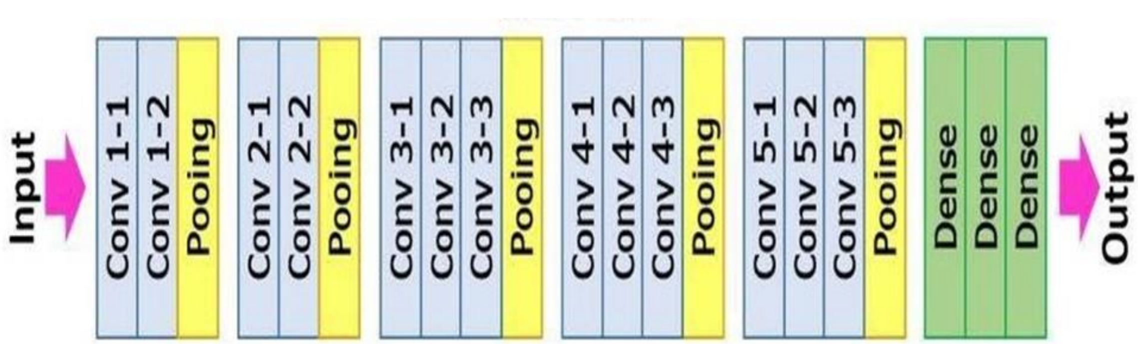

Fig 3.5: VGG16 Architecture [30]

3.3 Description of Fish Diseases

The development of an accurate classification system completely depends on in-depth disease study. The first step in this activity is image collection, during which many images are restored for use in training and testing. We have gathered images of both healthy and infected fish. The quality of images that are captured varies. The image is in RGB format.

3.3.1 Black Spot Disease

On the flesh, fins, and skin of fish, these parasitic organisms appear as small black spots. For the eradication of this issue, there is no appropriate control solution. The effect that this bacterium does on fish is minimal. The primary issue with black spots is the potentially ugly appearance they might produce. The majority of black spots can be removed by skinning diseased fish.

3.3.2 Red Spot Disease

Red Spot is a common fish disease. Red spot disease is also known as EUS. And it is caused by a fungus. Histological methods can quickly identify Red Spot in diseased fish samples procured from Red Spot-affected locations. Fish may have tiny sores or red patches. It isn't easy to control EUS in natural waters.

3.3.3 White Spot Disease

White Spot is the most well-known disease in fish. It is also known as ick or ich. The virus responsible for white spot syndrome is the cause of white spot disease. In addition to the spots on the fish, it also observes that they rub against tank-related objects to remove the spots. Loss of appetite may also be a symptom of the disease.

3.3.4 Healthy

Fish with spot-free bodies are considered to be healthy. Moreover, it will be disease-free.

3.4 Proposed Model

Our work enhances the proposed model in Figure 3.6. This model contains in mainly four phases. The first step is gathering the images. After which, the image was resized and enhanced in advance. Afterward, use five different pre-trained models. The ability to identify fish diseases will be our final achievement.

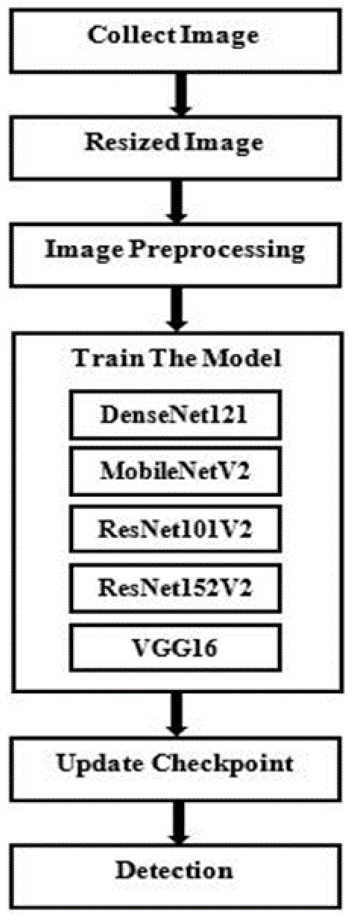

Fig 3.6: Proposed Model

3.5 Model Evaluation

The process of applying several evaluation metrics to understand how well a machine learning model performs is known as model evaluation. The efficiency of a model needs to be assessed early on in the research process, and model evaluation also helps with model monitoring. We can only evaluate our model based on training and validation accuracy. To assess our model, we must take into account several reports. To evaluate our model, we must create a classification report. The following subsections will discuss the brief description.

3.5.1 Classification Report

A classification report is a machine learning performance evaluation metric. The efficiency of classification algorithm prediction is evaluated using a classification report. The best model for this dataset cannot be determined only based on the accuracy score. In addition, we must assess a few factors that go into classification reports. The following details are provided:

3.5.1.1 Confusion Matrix

A performance measure table primarily depicts how well a machine learning model performed using a set of test output data. It calculates four terms to assess performance: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). In the experiment analysis portion, we will briefly explain this.

3.5.1.2 Accuracy

The number of samples from the total number of samples that were predicted correctly is the accuracy.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \text{ --------- } (1)$$

3.5.1.3 Precision

Precision is a ratio of True Positive results to all Positive predictions. This is sometimes referred to as positive predictive value.

$$\text{Precision} = TP / (TP + FP) \text{ --------- } (2)$$

3.5.1.4 Recall

It is the ratio between the number of accurate predictions and the true positive results. The recall is often referred to as sensitivity or true positive rate.

$$Recall = TP / (TP + FN) \text{ --------- } (3)$$

3.4.1.5 F1 Score

The F1 score is also called the harmonic mean of recall and precision.

$$F1\ Score = 2*TP / (2*TP + FP + FN) \text{ --------- } (4)$$

CHAPTER 4

Materials, Experiment Result and Discussion

4.1 Dataset

For the goal of our study, numerous sources of fish disease images and images of healthy fish were collected. Images of fish with black spot, white spot, red spot, and healthy fish are included in this collection. We chose datasets from the Kaggle dataset, GitHub repositories, and some useful websites. The collected datasets contain 315 images in total. Figure 4.1 displays examples of black spot, white spot, red spot, and healthy fish images.

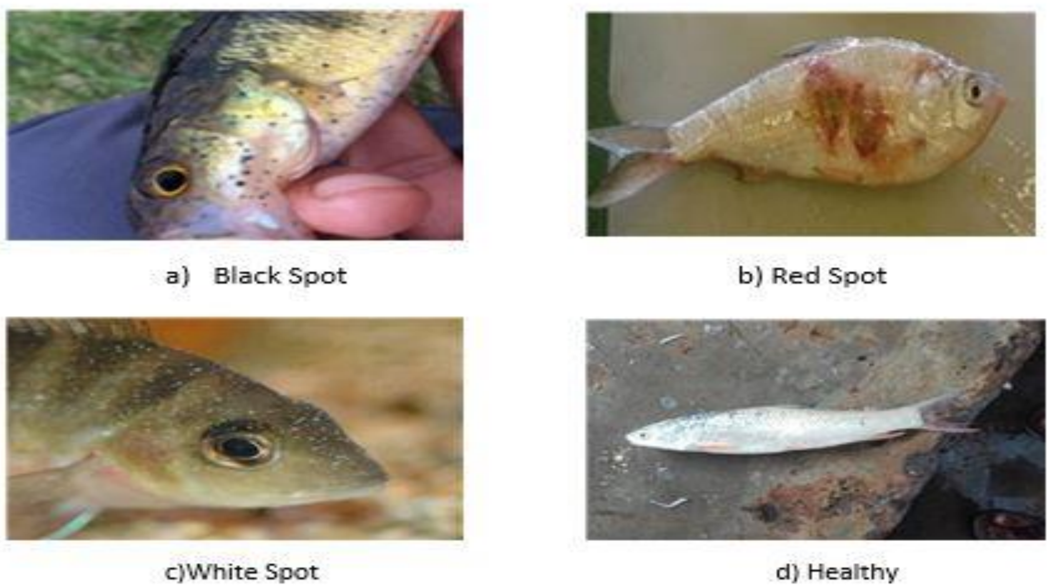

Fig 4.1: Fish images: a) Black spot, b) Red spot, c) White spot, d) Healthy.

4.2 Dataset Pre-Processing

To balance the four different types of datasets and increase the number of used images, the used datasets are first enhanced. According to the previous subsection, we have approximately 315 images, and after augmentation, we have a total of 1050 images. The four training classes will each employ augmentation techniques for the first period: black spot, white spot, red spot, and healthy. These images are then delivered into the data preprocessing stage.

Preparing the input data to meet the demands of the deep learning model is often done during the pre-processing stage. The input images for our model were prepared to utilize several pre-processing processes. Images are first scaled down to 224x224 pixels, and then they are augmented using techniques like flipping and rotating. The dataset needs to be divided in order to train the deep learning model. To ensure the variety of the images, the dataset images are divided into three sections (train, test, and validation). A total of 80% of the images are used for training, 10% are used for testing, and 10% are used for validation.

Table 4.1. The number of images per dataset

DATASET	BLACK SPOT IMAGES	RED SPOT IMAGES	WHITE SPOT IMAGES	HEALTHY IMAGES
TRAIN DATA	192	205	205	240
TEST DATA	24	25	25	30
VALIDATION DATA	24	25	25	30

4.3 Experiment Results and Analysis

Machine learning models can't always provide accurate predictions. Our models will try to provide the best solution. We used Python and the Keras framework to implement our models. Our proposed multi-classification deep learning models used the dataset we provided. So, we decided to use our dataset to train our models. An algorithm and appropriate fit methods were applied, with each model running for about 20 epochs and a batch size of 32, for training, testing, and validation of our models. We have to increase the number of samples and epochs to train deep-learning models effectively. We identified that the loss decreased when we used more epochs. Therefore, we employed 20 epochs. And, we use an automated detection approach with deep learning techniques. We run different deep-learning methods for detecting fish disease from images. Each of the architectures is thoroughly described in the study, and from the outcomes, it is determined which architecture is optimal and can reach the highest detection accuracy.

We observe that some algorithm accuracy has increased slightly after implementing our models based on our dataset, while other model accuracy has decreased. After running our models, the accuracy scores for our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models were 93.27%, 97.12%, 95.19%, 94.23%, and 96.15%, respectively.

In this project, different fishes are given as input to the model. The model is predicted for the above-given input and provides the output.

Fig 4.2: Experiment Results

The total amount of errors in our model is represented by the term loss. It assesses how well our model performed. If the errors are large, which suggests that the model is underperforming, the loss will also be considered. Otherwise, the lower it is, the better our model performs. Losses during the training and validation stages are also shown, together with the number of epochs, as in Figure 4.3. After applying our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models based on our dataset, we see a slight decrease in our algorithm loss values. The following loss values graphs for several pre-trained models:

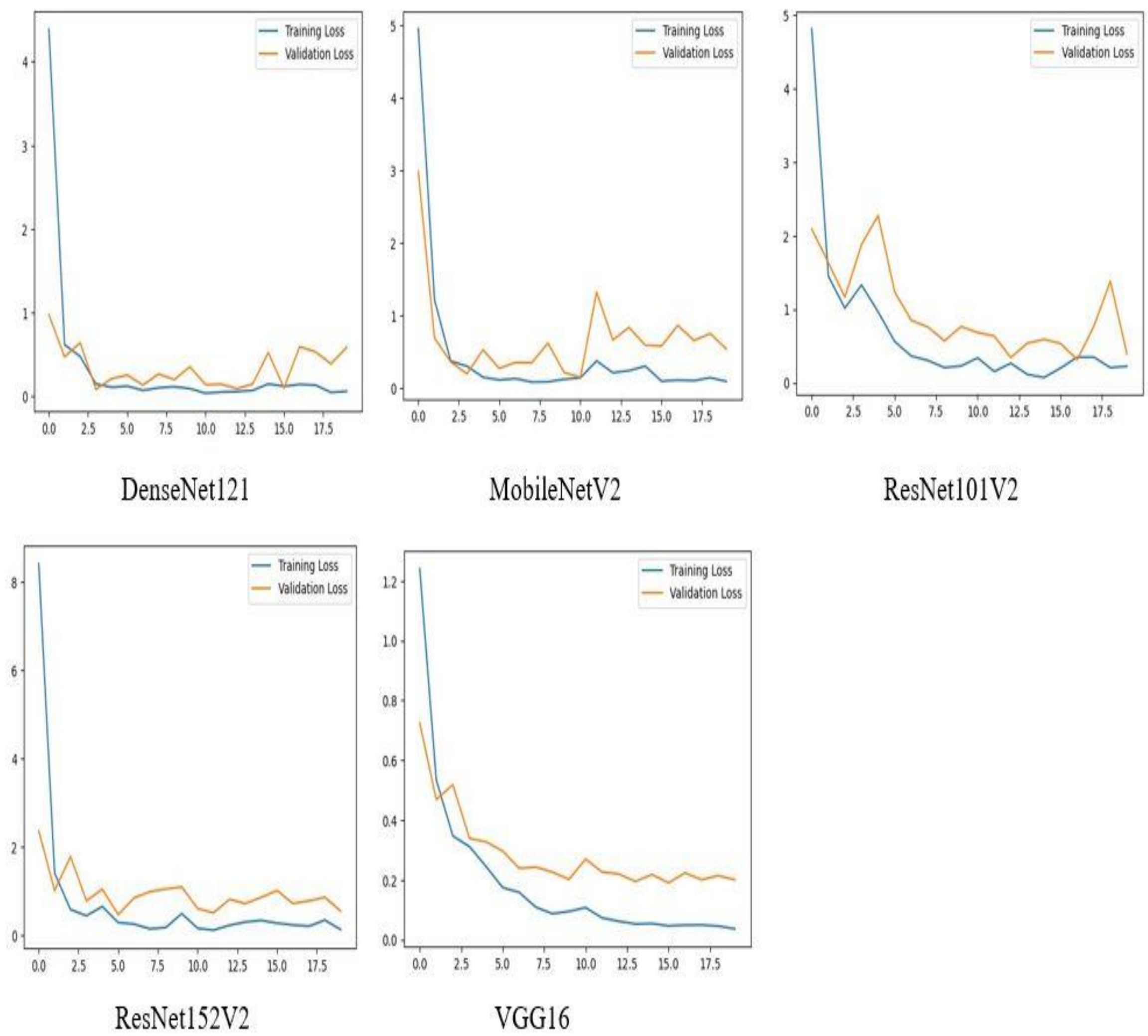

Fig 4.3: Curve of all model validation and training loss

The primary tool for assessing errors in multiclass classification is the confusion matrix. Using our dataset and all of the previously described pre-trained models, we first construct a confusion matrix. The visualization of the confusion matrix for DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models are shown in the figure. The confusion matrix for each algorithm we employed in our research is shown in Figure 4.4. The figure depicts how the MobileNetV2 model can classify the four classes (Black Spot, White Spot, Red Spot, and Healthy) with the highest ratio to the black spot images (24), followed by healthy (29), red spot (24), and white spot (24). Compared to the other proposed models for the four classes, this outcome ensures that the classification is done correctly.

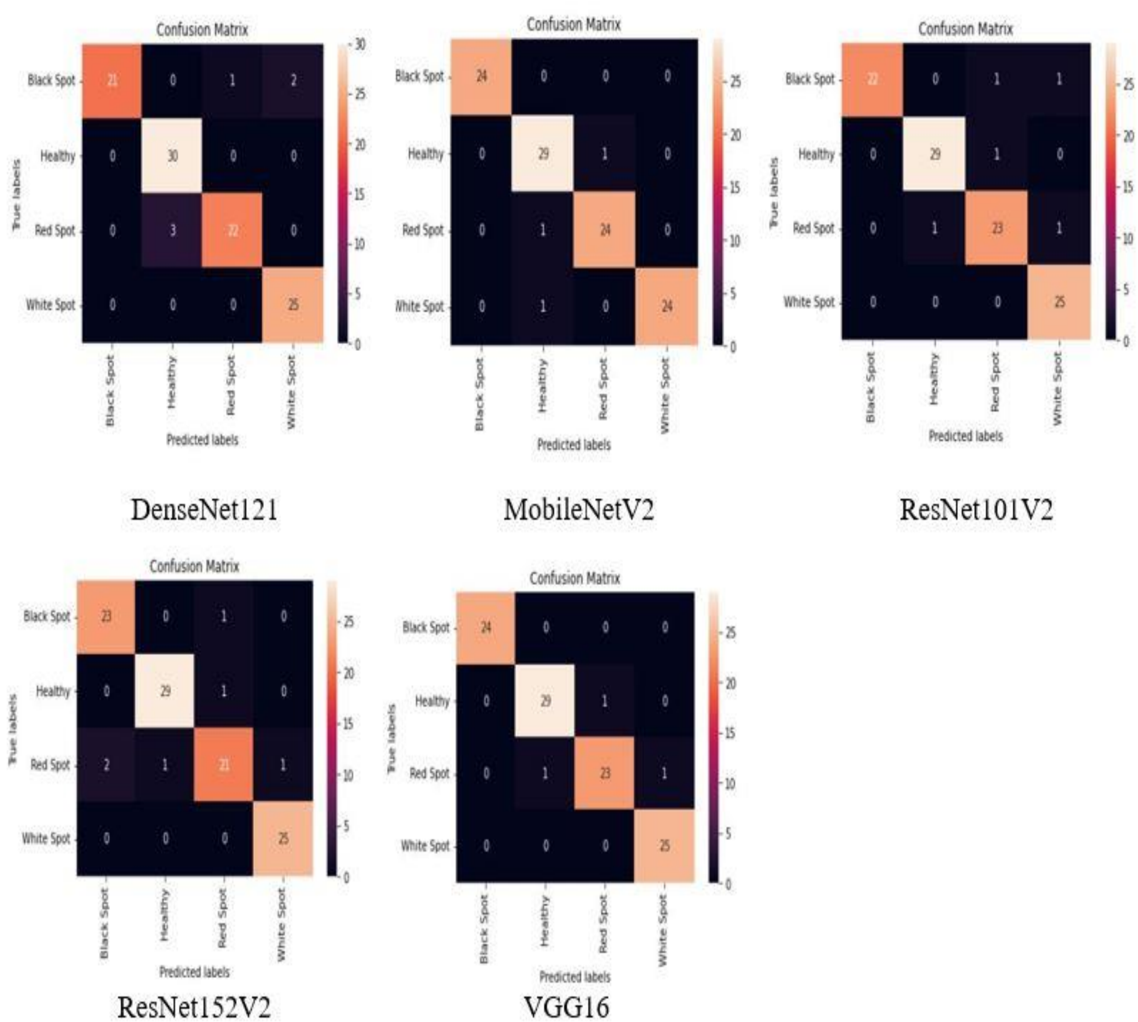

Figure 4.4: Confusion matrix of all models

The efficiency of predictions made by a classification algorithm is assessed using a classification report. We also assess our model using other metrics, including precision, recall, and f1-score. We create a classification report for our dataset using the confusion matrix. Using the confusion matrix, we determine the precision, recall, and f1-score. In terms of accuracy, precision, recall, and f1-score, the MobileNetV2 model performs better than the other four proposed models. This model achieved 93.27% accuracy, 94% precision, 93.27% recall, and 93.63% F1 score. The classification report for all of our models is displayed in Table 4.2.

Table 4.2: Classification Report

Algorithms Name	Accuracy	Precision	Recall	F1 Score
DenseNet121	93.27	94	93.27	93.63
MobileNetV2	97.12	97.17	97.12	97.15
ResNet101V2	95.19	95.33	95.19	95.26
ResNet152V2	94.23	94.17	94.23	94.2
VGG16	96.15	96.23	96.15	96.19

4.4 Different Model Results and Comparison

The experimental analysis of accuracy of the proposed models are shown in the figure 4.5. The DenseNet121 model attained an accuracy of 93.27%. The MobileNetV2 model achieved 97.12% accuracy. The ResNet101V2 model had an accuracy rate of 95.19%. An accuracy of 94.23% was attained with the ResNet152V2 model. The accuracy of the VGG16 model was 96.15%. DenseNet121 is the model that performs the diminutive overall, and MobileNetV2 is the model that performs the best overall.

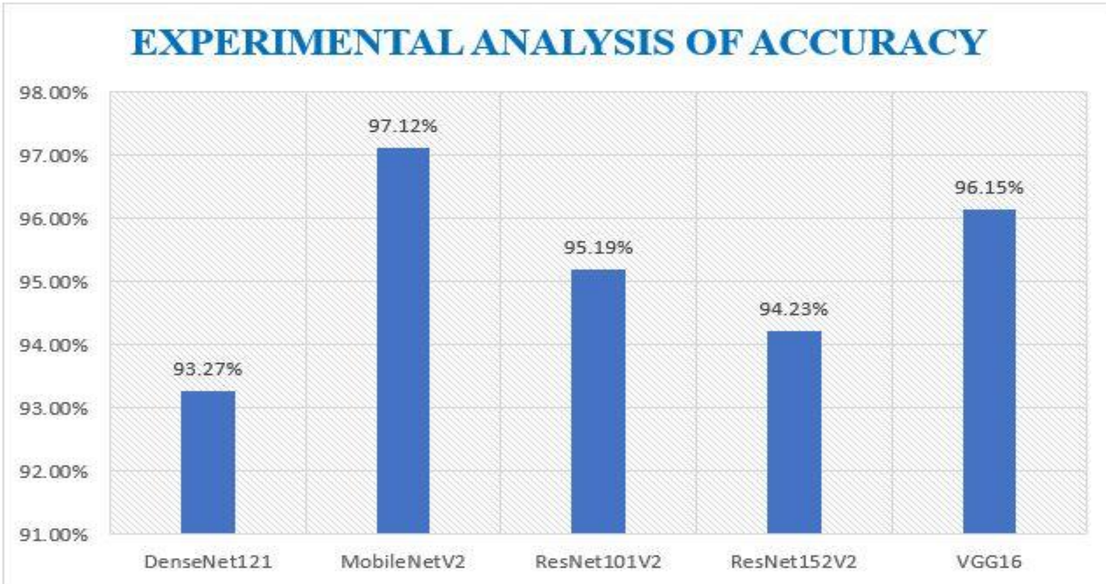

Fig 4.5: The accuracy of all proposed models

The experimental analysis of precision of the proposed models are shown in the figure 4.6. The precision scores for our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models were 94%, 97.17%, 95.33%, 94.17%, and 96.23%, respectively. The model that performs best all-around is MobileNetV2.

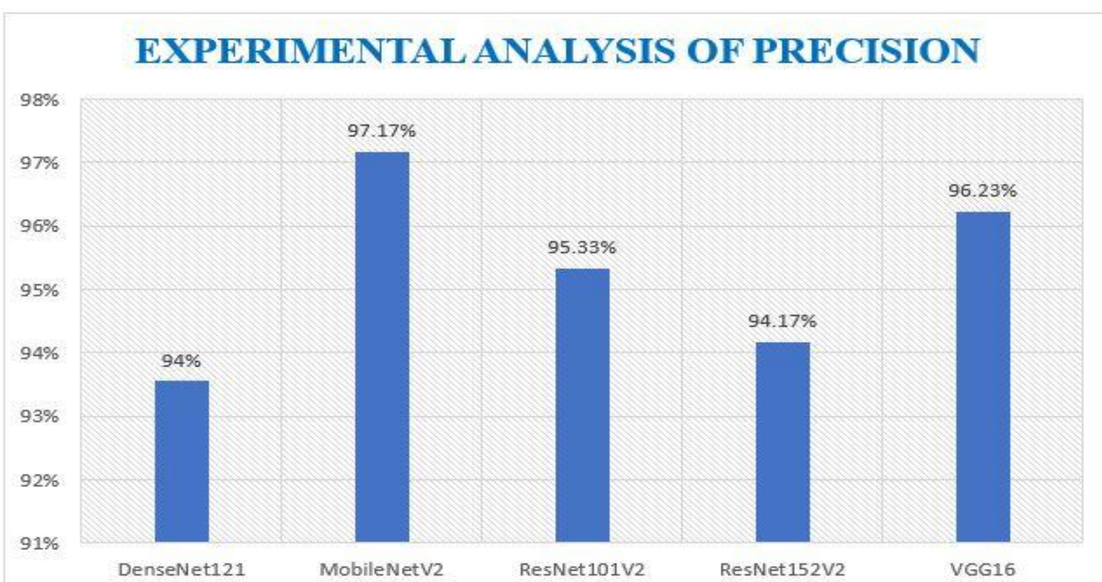

Fig 4.6: The precision of all proposed models

The experimental analysis of recall of the proposed models are shown in the figure 4.7. The recall scores for our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models were 93.27%, 97.12%, 95.19%, 94.23%, and 96.15%, respectively. The model that performs best all-around is MobileNetV2.

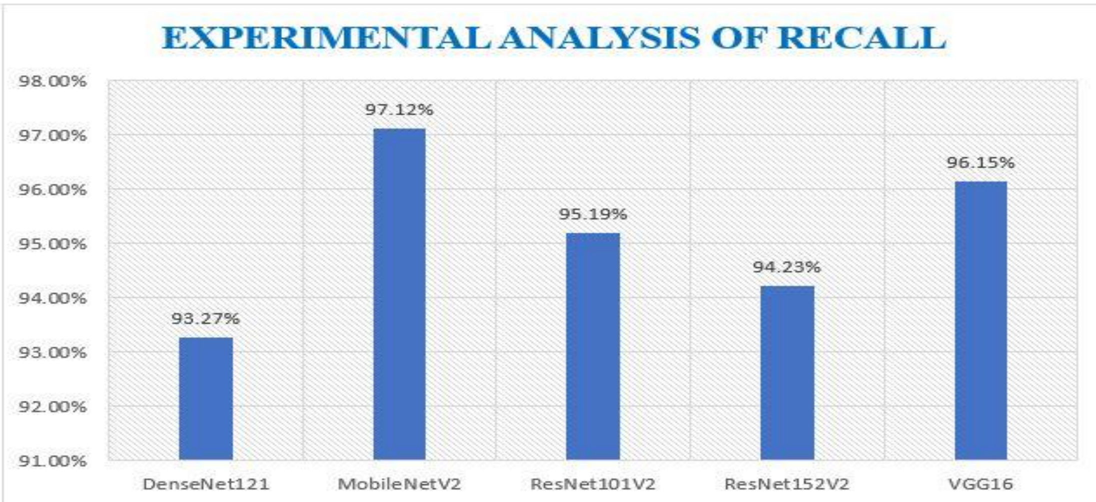

Fig 4.7: The recall of all proposed models

The experimental analysis of the f1-score of the proposed models are shown in the figure 4.8. The f1-score for our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models were 93.63%, 97.15%, 95.26%, 94.2%, and 96.19%, respectively. The model that performs best all-around is MobileNetV2.

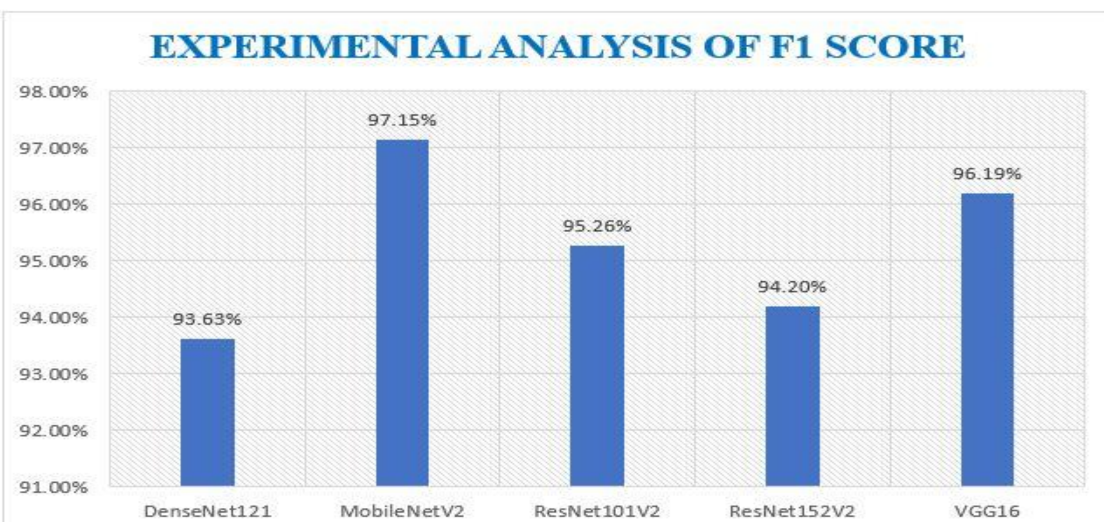

Fig 4.8: The f1-score of all proposed models

Figure 4.9 compares the models based on the various metrics: accuracy, precision, recall, and F1-score for the proposed models. The MobileNetV2 model exceeds the other four proposed models in terms of accuracy, precision, recall, and F1-score, as is evident from the figures. The MobileNetV2 model achieved 93.27% accuracy, 94% precision, 93.27% recall, and 93.63% F1 score.

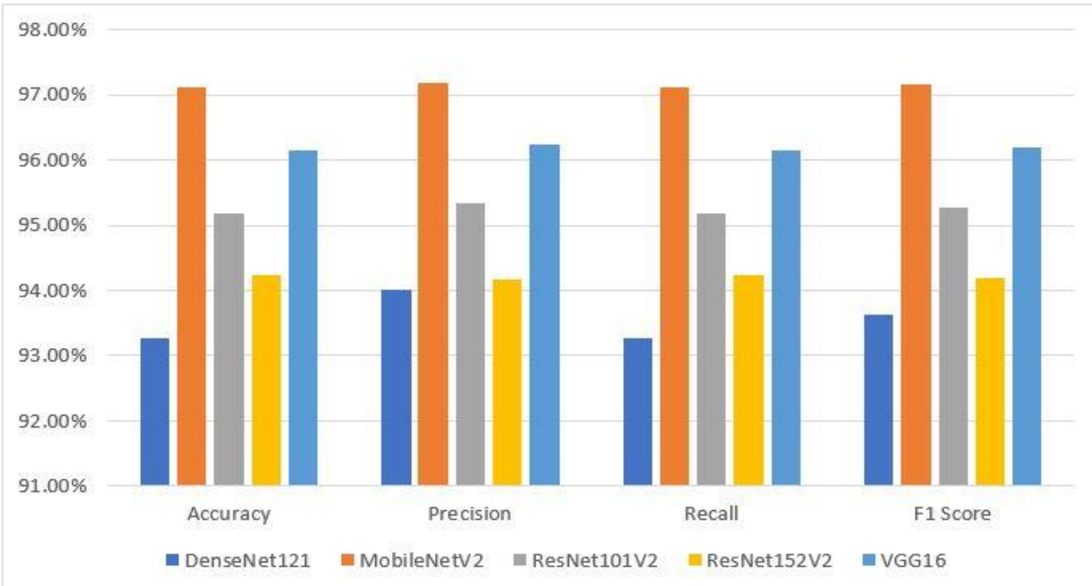

Fig 4.9: Accuracy, precision, recall, and f1-score for the proposed models

4.5 Discussion

The system for fish disease identification in the current work is proposed with five deep-learning architectures. Black spots, red spots, and white spots are the most common fish diseases, and these architectures are used to identify them. We display our project's step by-step operation and the accuracy of the output in the experiment results and result from analysis. After running our project, the accuracy scores for our DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 models were 93.27%, 97.12%, 95.19%, 94.23%, and 96.15%, respectively. According to the experiment data, the MobileNetV2 model performed better than the other proposed models. In comparison to other models, the model provided good detection accuracy.

CHAPTER 5

Impact on Society, Environment and Sustainability

5.1 Impact on Society

Fish disease is a major issue that is steadily harming our society. If we do nothing immediately to address the serious issue of fish disease, it could lead to poor fish production or low-level mortalities, limiting productivity. We need to increase knowledge among ourselves without relying on the government since we are aware that farmers will not identify the fish disease properly is the main contributor to its development. To resolve this issue, we proposed a model elevated by society, and we hope our research will benefit the fish industry.

5.2 Impact on Environment

We believe our technology has a positive impact on the environment as well. Fish with diseases that survive may be difficult to sell. The productivity and financial losses brought on by fish disease make it challenging for farmers to support themselves. Through the use of our technology, fish industry professionals will be able to detect fish diseases.

5.3 Ethical Aspects

Our model is completely open-source and free, which benefits the usual people. These are the main ethical principles we adhere to. That is why we take a precaution to protect our fish industry. Regarding our model, we also have some ethical aspects:

- To prevent fish disease.
- To protect our fish industry.
- Fish farmers should be increased their knowledge.

5.4 Sustainability Plan

Our goal is to help the fish industry and increase fish production. Although this algorithm is free, it requires funding from the government or investors to implement a distinct industry strategy. If this method is included in the nation's fish sector, there is a significant chance that the project will succeed.

CHAPTER 6

Overview of the Study, Conclusion and Future Work

6.1 Overview of the Study

Deep learning is a vast area that can identify everything. In this analysis, we showed the way to fish disease identification. We develop a model that can identify fish disease promptly. We accomplish this by employing deep learning methods for automatic detection. We run different deep-learning techniques for detecting fish disease from images. The project has six steps: collect image, resize image, image preprocessing, train, update checkpoint, and detection. Moreover, we compare deep-learning architectures using datasets for four classes of fish diseases.

6.2 Conclusion

In this study, a multi-classification deep learning model was evaluated and developed for fish disease identification. Even though identifying fish diseases is difficult, we did our job perfectly. This deep learning model with many classifications can identify healthy, black, white, and red spots from fish images. This model employs various deep-learning methods to identify fish disease from images. In this work, the architectures DenseNet121, MobileNetV2, ResNet101V2, ResNet152V2, and VGG16 were all given. The MobileNetV2 model outperformed the other four suggested models in thorough tests and results on datasets gathered from various sources.

6.3 Future Work

In this study, the algorithm that we used currently has several limitations.

Limitation:

- The experiment has a very high success rate for better accuracy identification.
- Efficiency may remain the same with a larger data collection or decrease slightly.
- It needs a highly configurable GPU processor.

Future Scope:

- We will use other deep-learning techniques to develop our project.
- We will try to increase the training epochs and the number of images in the used datasets.
- To reduce expenses, we'll try to make it into a small device.

References

[1] M. M. Shamsuzzaman, M. M. Islam, N. J. Tania, M. A. Al-Mamun, P. P. Barman and X. Xu, "Fisheries resources of Bangladesh: Present status and future direction," Aquaculture and Fisheries, vol. 2, no. 4, pp. 145-156, 2017.

[2] 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.

[3] J. Sikder, K. I. Sarek and U. K. Das, "Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh," (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 12, no. 6, pp. 867- 871, 2021.

[4] S. Malik, T. Kumar and A. K. Sahoo, "Image Processing Techniques for Identification of Fish Disease," in IEEE 2nd International Conference on Signal and Image Processing, 2017.

[5] N. Hasan, S. Ibrahim and A. A. Azlan, "Fish diseases detection using convolutional neural network (CNN)," International Journal of Nonlinear Analysis and Applications, vol. 13, no. 1, pp. 1977-1984, 2022.

[6] J.-S. Park, M.-J. Oh and S. Han, "Fish Disease Diagnosis System Based on Image Processing of Pathogens' Microscopic Images," in Frontiers in the Convergence of Bioscience and Information Technologies, 2007.

[7] S. N. Pauzi, M. G. Hassan, N. Yusoff, N. H. Harun, A. H. A. Bakar and B. C. Kua, "A review on image processing for fish disease detection," Journal of Physics: Conference Series, vol. 1997, no. 1, p. 012042, 2021.

[8] B. Austin, "Methods for the diagnosis of bacterial fish diseases," Marine Life Science & Technology, vol. 1, no. 1, pp. 41-49, 2019.

[9] H. Charo-Karisa, S. E. Ali, E. Marijani, N. A. Ibrahim, T. Q. Trinh, M. V. Chadag and J. A. Benzie, "Genetic parameters for black spot disease (diplopstomiasis) caused by Uvulifer sp. infection in Nile tilapia (Oreochromis niloticus L.)," Aquaculture, vol. 532, p. 736039, 2021.

[10] A.-A. Nayan, J. Saha, A. N. Mozumder, 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," TRENDS IN SCIENCES, vol. 18, no. 21, pp. 351-351, 2021.

[11] M. J. Mia, R. B. Mahmud, M. S. Sadad, H. A. Asad and R. Hossain, "An in-depth automated approach for fish disease recognition," Journal of King Saud University-Computer and Information Sciences, pp. 1-10, 2022.

[12] R. Saeed, H. Feng, X. Wang, X. Zhang and Z. Fu , "Fish quality evaluation by sensor and machine learning: A mechanistic review," Food Control, vol. 137, p. 108902, 2022.

[13] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L. R. Goulart, C. E. Davis and A. M. Dandekar, "Advanced methods of plant disease detection. A review," Agronomy for Sustainable Development, vol. 35, no. 1, pp. 1-25, 2015.

[14] S. P. Mohanty, D. P. Hughes and M. Salathe, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in plant science, vol. 7, p. 1419, 2016.

[15] S. Ramesh, M. Ramachandra Hebbar, N. M, P. R, P. Bhat N, S. N and M. P V Vinod, "Plant Disease Detection Using Machine Learning," in 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C), 2018.

[16] V. B. Kumar, S. S Kumar and V. Saboo, "Dermatological Disease Detection Using Image Processing and Machine Learning," in 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), 2016.

[17] N. S. A. ALEnezi, "A Method Of Skin Disease Detection Using Image Processing And Machine Learning," Procedia Computer Science, vol. 163, pp. 85-92, 2019.

[18] K. Ahmed, T. R. Shahidi, S. M. I. Alam and S. Momen, "Rice Leaf Disease Detection Using Machine Learning Techniques," in 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019.

[19] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong and C. Zhou, "Deep learning for smart fish farming: applications, opportunities and challenges," Reviews in Aquaculture, vol. 13, no. 1, pp. 66-90, 2021.

[20] T. H. Khai, S. N. H. S. Abdullah, M. K. Hasan and A. Tarmizi , "Underwater Fish Detection and Counting Using Mask Regional Convolutional Neural Network," Water, vol. 14, no. 2, p. 222, 2022.

[21] V. P and J. L. R. J, "Recognition Of Fish Categories Using Deep Learning Technique," in 3rd International Conference on Computing and Communication Technologies ICCCT 2019, 2019.

[22] 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.

[23] A. N.S., S. D. and R. K. S., "Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries," Ecological Informatics, vol. 61, p. 101248, 2021.

[24] E. Yavuzer and M. Kose, "Prediction of fish quality level with machine learning," International Journal of Food Science and Technology, vol. 57, no. 8, pp. 5250-5255, 2022.

[25] L. Yang, Y. Liu, H. Yu, X. Fang, L. Song, D. Li and Y. Chen, "Computer Vision Models in Intelligent Aquaculture with Emphasis on Fish Detection and Behavior Analysis: A Review," Archives of Computational Methods in Engineering, vol. 28, no. 4, pp. 2785-2816, 2021.

[26] ResearchGate, available at << https://www.researchgate.net/figure/Left-DenseNet121-architecture-Right-Dense-block-conv-block-and-transition-layer_fig4_331364877>>, last accessed on 20-11-2022 at 11:00 PM.

[27] Data Driven Investor, available at << https://medium.datadriveninvestor.com/review-on-mobile-net-v2-ec5cb7946784>>, last accessed on 20-11-2022 at 11:20 PM.

[28] ResearchGate, available at << https://www.researchgate.net/figure/Original-architecture-of-ResNet-101-deep-learning-model_fig3_354871169>>, last accessed on 23-11-2022 at 7:00 PM.

[29] ResearchGate, available at << https://www.researchgate.net/figure/ResNet101V2-and-ResNet152V2-architecture-with-extra-layers-added-at-the-end-for_fig4_353225711>>, last accessed on 23-11-2022 at 7:30 PM.

[30] Great Learning, available at << https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918 >>, last accessed on 23-11-2022 at 8:20 PM.

Plagiarism Report

A COMPARATIVE ANALYSIS OF DEEP LEARNING APPROACHES FOR FISH DISEASE IDENTIFICATION

ORIGINALITY REPORT

16%	13%	7%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
2	Submitted to Daffodil International University Student Paper	2%
3	ijnaa.semnan.ac.ir Internet Source	1%
4	www.researchgate.net Internet Source	1%
5	"Trends in Data Engineering Methods for Intelligent Systems", Springer Science and Business Media LLC, 2021 Publication	1%
6	bio-protocol.org Internet Source	1%
7	www.mdpi.com Internet Source	1%
8	"SCSE 2022 Conference Proceedings", 2022 International Research Conference on Smart	<1%