FISH DISEASE DETECTION USING IMAGE BASED MACHINE LEARNING TECHNIQUES IN AQUACULTURE

K Venkatesh[1], Nandamuri Veda Kumari[2]

[1] **Assistant Professor MCA, DEPT,** Dantuluri Narayana Raju College **, Bhimavaram, Andharapradesh**
Email id:- kornalavenkatesh@gmail.com
[2]**PG Student of MCA,** Dantuluri Narayana Raju College **, Bhimavaram, Andharapradesh**
Email id :- vedanandamuri123@gmail.com

ABSTRACT

Fish diseases in aquaculture constitute a significant hazard to nutriment security. Identification of infected fishes in aquaculture remains challenging to find out at the early stage due to the dearth of necessary infrastructure. The identification of infected fish timely is an obligatory step to thwart from spreading disease. Inthis work, we want to find out the salmon fish disease in aquaculture, as salmon aquaculture is the fastest-growing food production system globally, accounting for 70 percent (2.5 million tons) of the market. In the alliance of flawless image processing and machine learning mechanism, we identify the infected fishes caused by the various pathogen.

This work divides into two portions. In the rudimentary portion, image pre-processing and segmentation have been applied to reduce noise and exaggerate the image, respectively. In the second portion, we extract the involved features to classify the diseases with the help of the Support Vector Machine (SVM) algorithm of machine learning with a kernel function. The processed images of the first portion have passed through this (SVM) model. Then we harmonize a comprehensive experiment with the proposed combination of techniques on the salmon fish image dataset used to examine the fish disease. We have conveyed this work on a novel dataset compromising with and without image augmentation. The results have bought a judgment of our applied SVM performs notably with 91.42 and 94.12 percent of accuracy, respectively, with and without augmentation.

1. INTRODUCTION

Indian fisheries and aquaculture is an important sector of food production, providing nutritional security to the food basket, contributing to the agricultural exports and engaging about 14 million people in different activities. Aquaculture resources in India include 2.36 million ha of ponds and tanks, 0.798 million ha of flood plain lakes plus in addition 195 210 km of rivers and canals, 2.907 million of reservoirs and that could be utilized for aquaculture purposes. Ponds and tanks are the prime resources for freshwater aquaculture; however, only about 40 percent of the available area is used for aquaculture currently. Freshwater aquaculture activity being an important activity expanded its dimension in terms of area coverage and intensity of operation with Andhra Pradesh, Punjab, Haryana, Maharashtra etc, taking up fish culture as a commercial farming enterprise. Of late, scientific carp farming is picking up in the north-eastern states of India. Brackish water aquaculture is mainly concentrated on the coasts of Andhra Pradesh, Tamil Nadu, Orissa and West Bengal. With regards to the market, while the main areas of consumption for freshwater fish are in West Bengal, Bihar, Orissa and north-eastern India. Cultured brackish water

shrimps are destined mainly for export.

While carp form the most important species farmed in freshwater in India, it is the shrimp from the brackish water sector which contributes the bulk of the production. The three Indian majorcarps, namely, catla (Catla catla), rohu labeo (Labeo rohita) and mrigal carp (Cirrhinus mrigala) contribute over 90% of the total Indian aquaculture production. Introduced during the 1970's into the carp polyculture system in the country, three exotic carps, namely, silver carp (Hypophthalmichthys molitrix), grass carp (Ctenopharyngodon idellus) and common carp (Cyprinus carpio) now form a second important group.

2. LITERATURE SURVEY AND RELATED WORK

Start by researching different machine learning algorithms commonly used for image analysis, such as convolutional neural networks (CNNs) and support vector machines (SVMs).

1. Look for studies that have applied these algorithms to detect specific fish diseases, such as bacterial infections, parasitic infestations, or viral outbreaks.

2. Explore the datasets used in these studies, including the types of fish species, the number of images, and the annotation methods employed.

3. Pay attention to the image preprocessing techniques used to enhance the quality of the images and improve the accuracy of disease detection.

4. Examine the performance metrics used to evaluate the effectiveness of the machine learning models, such as accuracy, precision, recall, and F1-score.

5. Consider the challenges and limitations faced in these studies, such as limited dataset availability, variations in fish appearance, and the need for real-time detection in aquaculture settings.

3. EXISTING SYSTEM

Some works focused on only some basic image processing techniques for the identification of fish disease. Shaveta et al. [22] proposed an image-based detection technique where firstly applies image segmentation as an edge detection with Canny, Prewitt, and Sobel. However, they did not specify the exact technique that engrossed for feature extraction. In feature extraction, they applied Histogram of Gradient (HOG) and Features from Accelerated Segment Test (FAST) for classification with a combination of both techniques. They tried to discover a better classification with a combination instead of applying a specific method with less exactness. Another technique Lyubchenko et al. [21] proposed a structure called the clustering of objects in the image that obliged diverse image segmentation actions based on a scale of various clusters. Here, they chose markers for individual objects and objects encountered with a specific marker. Finally, they calculated the proportion of an object in the image and the proportion of infected area to the fish body to identify fish disease. However, individual marking of an object is time consuming and not effective.

There are some approaches focused on the consolidation of image processing and machine learning. Malik et al. [23] proposed a

specific fish disease called Epizootic Ulcerative Syndrome (EUS) detection approach. Aphanomyces invadans, a fungal pathogen, cause this disease. Here, they approached combination styles that combine the Principal Component Analysis (PCA) and Histogram of Oriented Gradients (HOG) with Features from Accelerated Segment Test (FAST) feature detector and then classify over machine learning algorithm (neural network). The sequence of FAST-PCANN gives 86 percent accuracy through the classifier, and HOG-PCA-NN gives 65.8 percent accuracy that is less than the previous combination.

DISADVANTAGES:

1.Fish diseases in aquaculture constitute a significant hazard to nutriment security.

2.The identification of infected fish timely is an obligatory step to thwart from spreading disease. In this work, we want to find out the salmon fish disease in aquaculture, as salmon aquaculture is the fastest-growing food production system globally, accounting for 70 percent (2.5 million tons) of the market.

3. In the alliance of flawless image processing and machine learning mechanism, we identify the infected fishes caused by the various pathogen.

4. PROPOSED SYSTEM

One of the most popular supervised machine learning techniques, support vector machine (SVM), has brought convenient solutions for many classification problems in various fields. It is a powerful classification tool that brings out quality predictions for unlabeled data. In [19] Authors built an SVM model based on three kernel functions to differentiate dengue human infected blood sera and healthy sera. For image classification, another SVM architecture has been proposed in [3] where they emulate the architecture by combining convolutional neural network (CNN) with SVM. SVM provides remarkable accuracy in many contexts.

ADVANTAGES:

1. Fish diseases in aquaculture constitute a significant hazard to nutriment security.
2. Identification of infected fishes in aquaculture remains challenging to find out at the early stage due to the dearth of necessary infrastructure.
3. The identification of infected fish timely is an obligatory step to thwart from spreading disease. In this work, we want to find out the salmon fish disease in aquaculture, as salmon aquaculture is the fastest-growing food production system globally, accounting for 70 percent (2.5 million tons) of the market.

6. RESULTS AND DISCUSSION SCREEN SHOTS

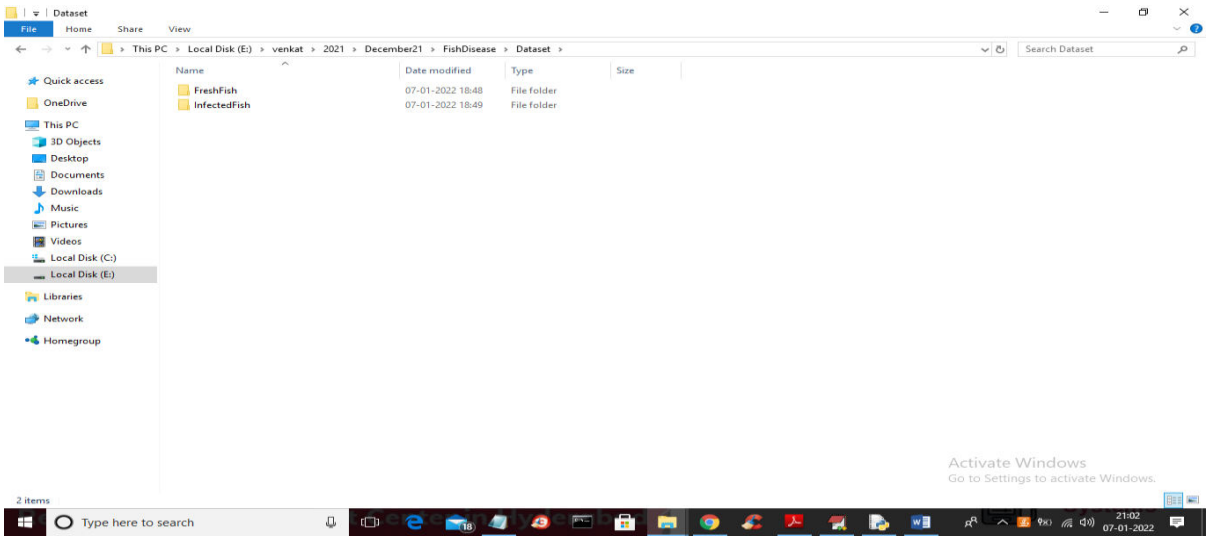
Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture Worldwide 20% of the revenue generated from Aquaculture (sea food) and it's important to produce and deliver fresh sea food and to identify INFECTED or FRESH fish's author of this paper is using various machine learning algorithm such as SVM, Decision Tree, Naïve Bayes and Logistic Regression and among all algorithm SVM is giving better performance.

Before training ML model author has applied various preprocessing task on images such as Cubic Splines Interpolation, CLAHE and LAB KMEANS colour segmentation

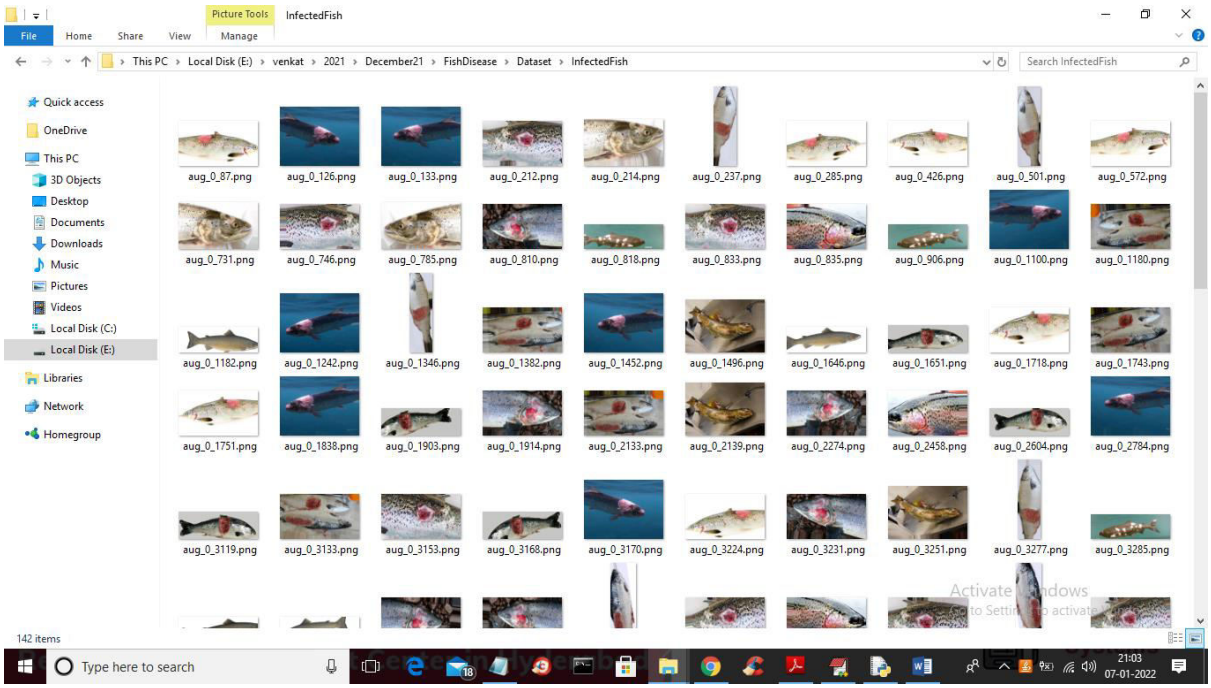
1) Cubic Splines Interpolation: using this module we will apply image magnification and fixed-size conversion, we use an improved interpolation method called extended Cubic Splines interpolation
2) Adaptive Histogram Equalization: using this module we will enhance image contrast by applying histogram algorithm called CLAHE (contrast limited adaptive histogram equalization)
3) RGB Colour Space to L*a*b Colour Space: using this module we will segment image by using colour format. LAB algorithm will use KMEANS algorithm to segment similar colours into same cluster.

We will process all images using above 3 techniques and then processed images will be input to SVM algorithm to build fish freshness prediction model.

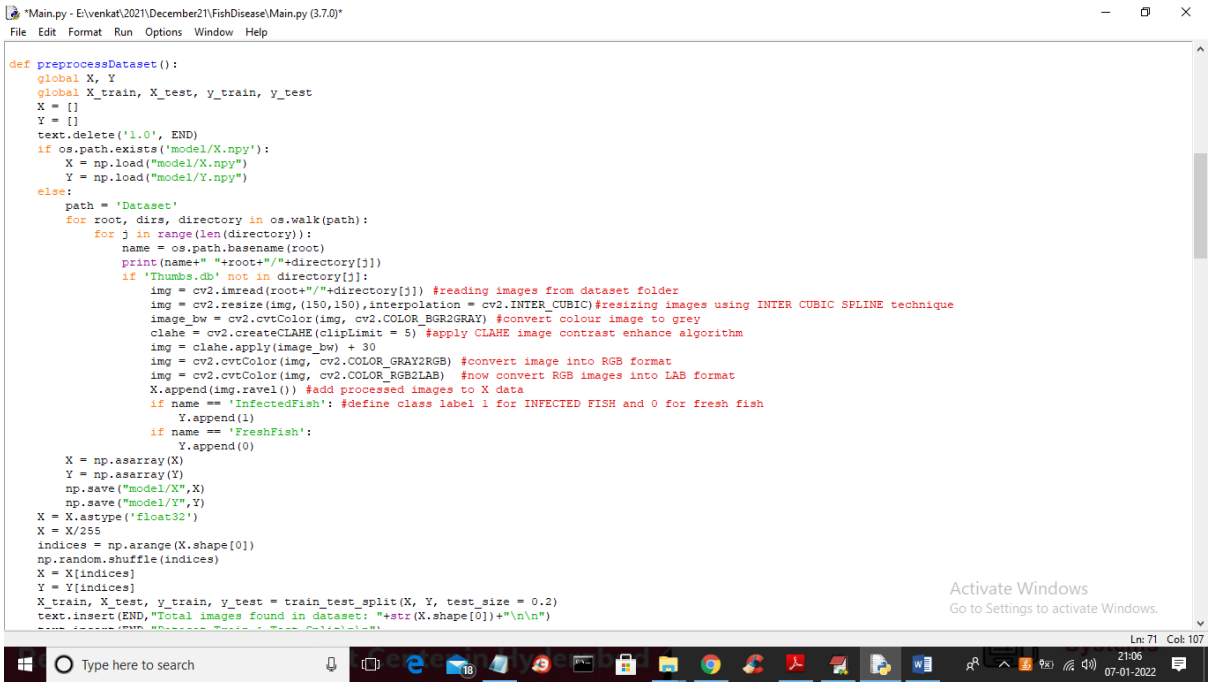
To implement this project, we have used below images



In above screen we have two folders called 'Fresh Fish" and "Infected Fish" inside dataset folder and you can go inside any folder to see either fresh or infected fish images



In above screen we can see all infected fish images. In below screen we can see code for SPLINE CUBIC, CLAHE and LAB implementation



In above screen read red colour comments to know about coding

To implement this project we have designed following modules

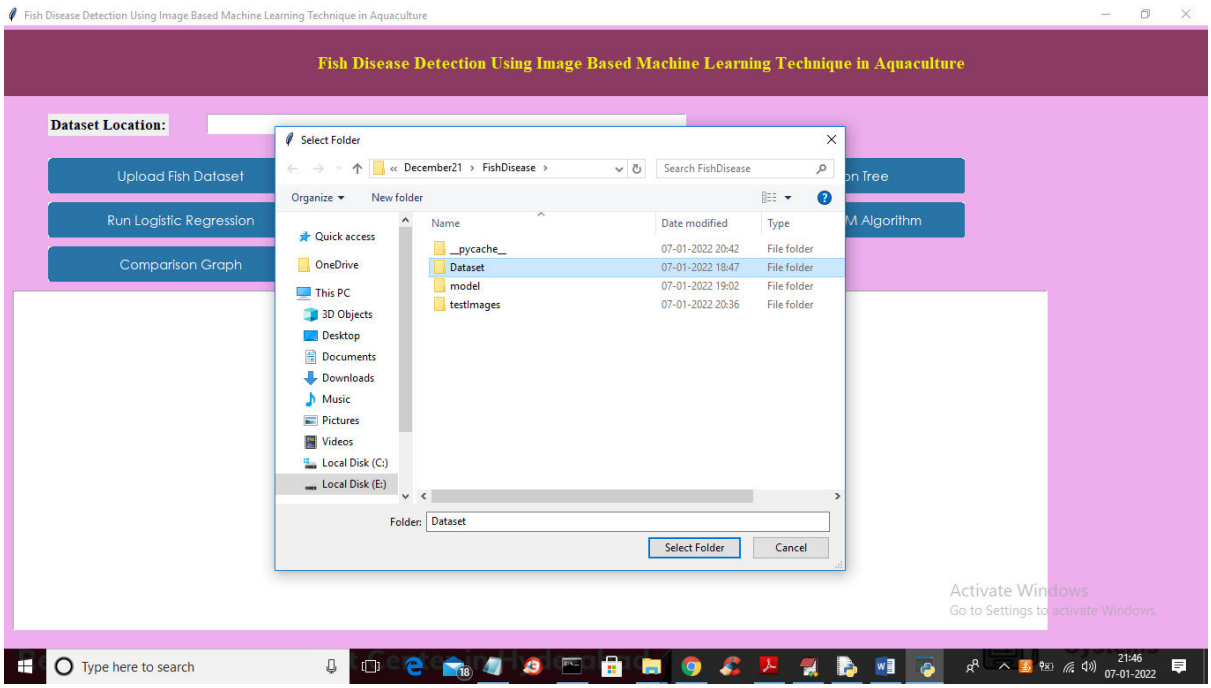
1) Upload Fish Dataset: using this module we will upload dataset to application
2) Run Interpolation, CLAHE & LAB: using this module we will read all images and then apply interpolation, CLAHE and LAB to process all images and then normalize images and then split dataset into train and test

3) Run Decision Tree: processed train images will be input to decision tree to trained a model and this model will be applied on TEST images to calculate prediction accuracy and other metrics
4) Run Logistic Regression: processed train images will be input to logistic regression to trained a model and this model will be applied on TEST images to calculate prediction accuracy and other metrics
5) Run Naive Bayes: processed train images will be input to naïve bayes to trained a model and this model will be applied on TEST images to calculate prediction accuracy and other metrics
6) Run Propose SVM Algorithm: processed train images will be input to SVM algorithm to trained a model and this model will be applied on TEST images to calculate prediction accuracy and other metrics
7) Comparison Graph: using this module we will plot accuracy and other metric graphs
8) Predict Fish Status: using this module we will upload test image and then SVM algorithm will predict whether image contains fresh or infected fish.

SCREEN SHOTS:

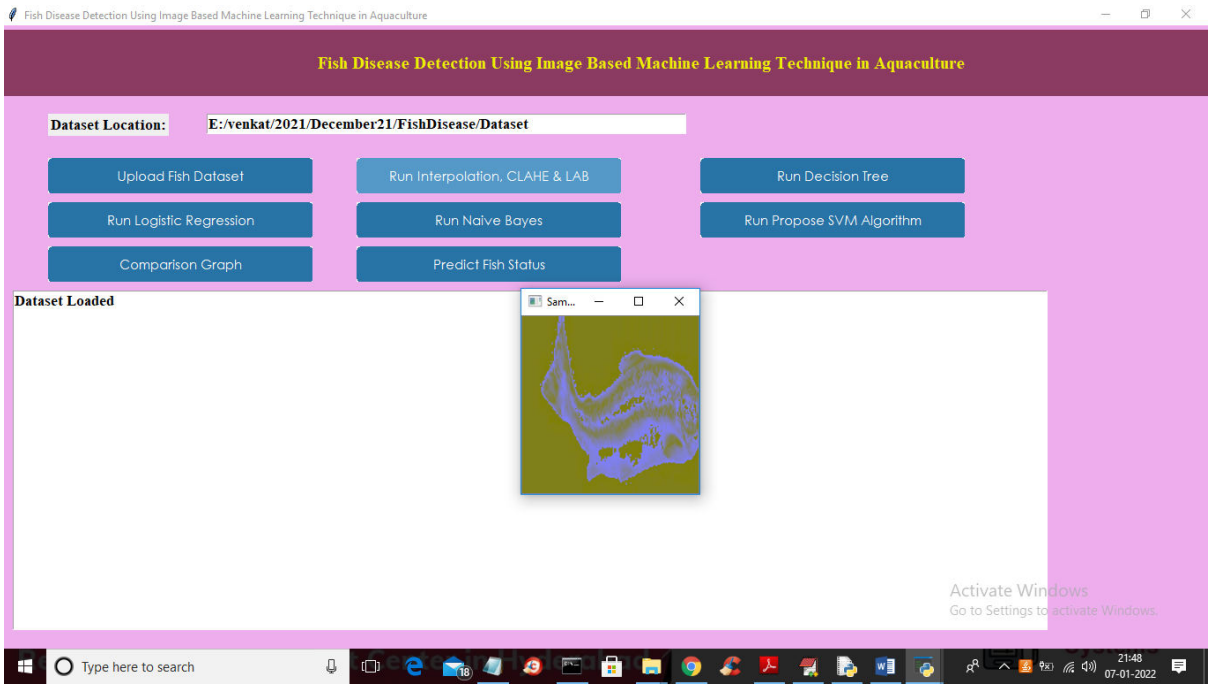
To run project double click on 'run.bat' file to get below screen

In above screen click on 'Upload Fish Dataset' button to upload dataset and to get below screen



In above screen selecting and uploading 'Dataset' folder and then click on 'Select Folder' button to load dataset and to get below screen

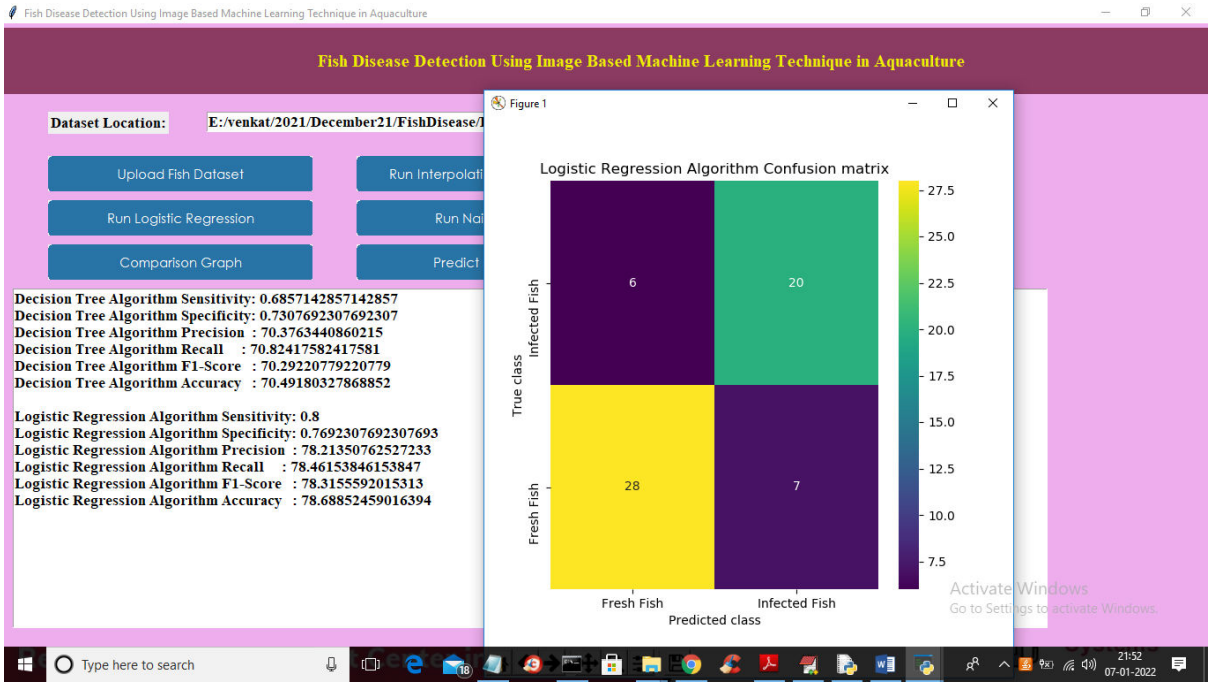
In above screen dataset loaded and now click on 'Run Interpolation, CLAHE & LAB' button to apply all 3 techniques and to process images and then split dataset into train and test



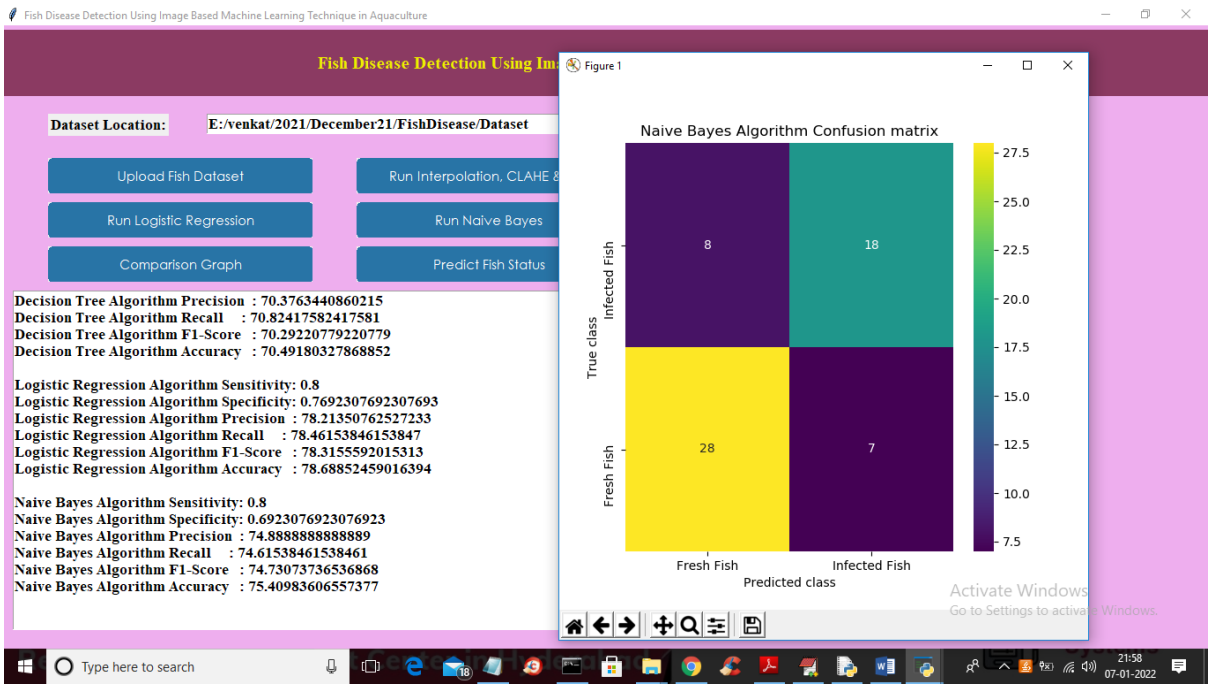
In above screen we can see all images are process using 3 techniques and after processing image will be converted to above format and now close above image to get below screen

In above screen we can see dataset contains 305 image and then split into 80 and 20% train and test data to get 244 as training images and 61 as test images. Now process images are ready with train and test data and now click on 'Run Decision Tree' button to train decision tree and to get below screen

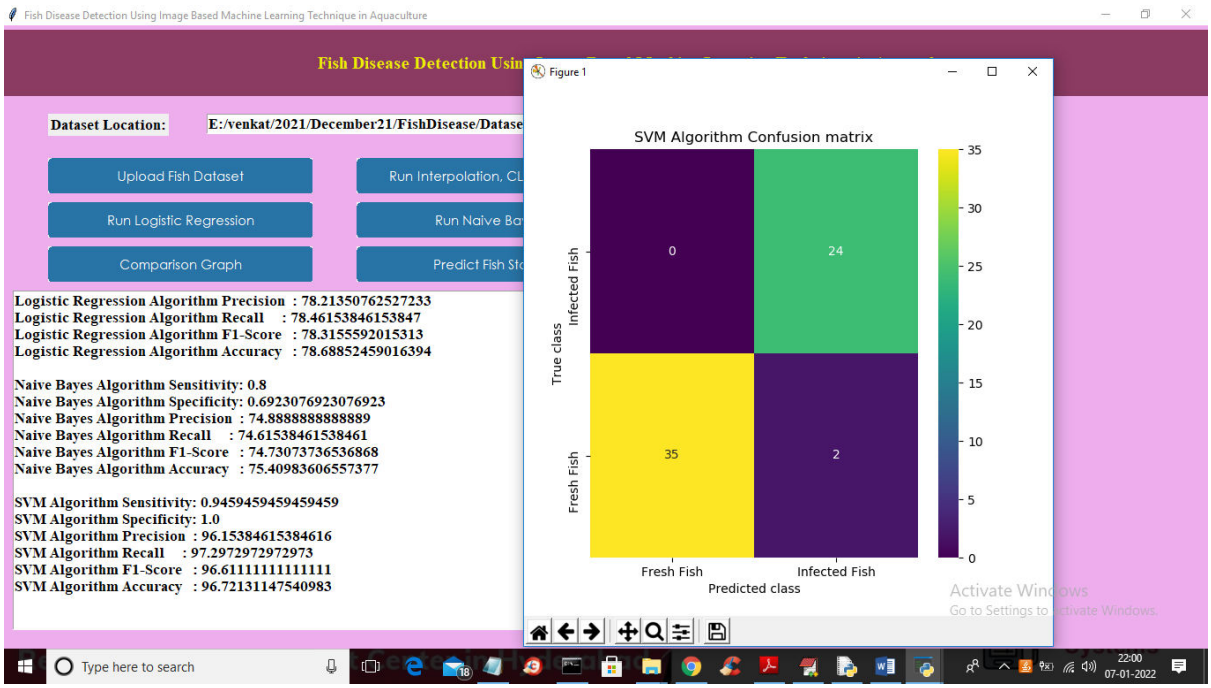
In above screen with decision tree we got 70% accuracy and we can see other metric values also and in confusion matrix graph we can see 24 images are correctly predicted as Fresh fish and 7 is wrongly predicted as infected fish. Now close above graph and then click on 'Run Logistic Regression' button to train logistic regression algorithm and to get below output



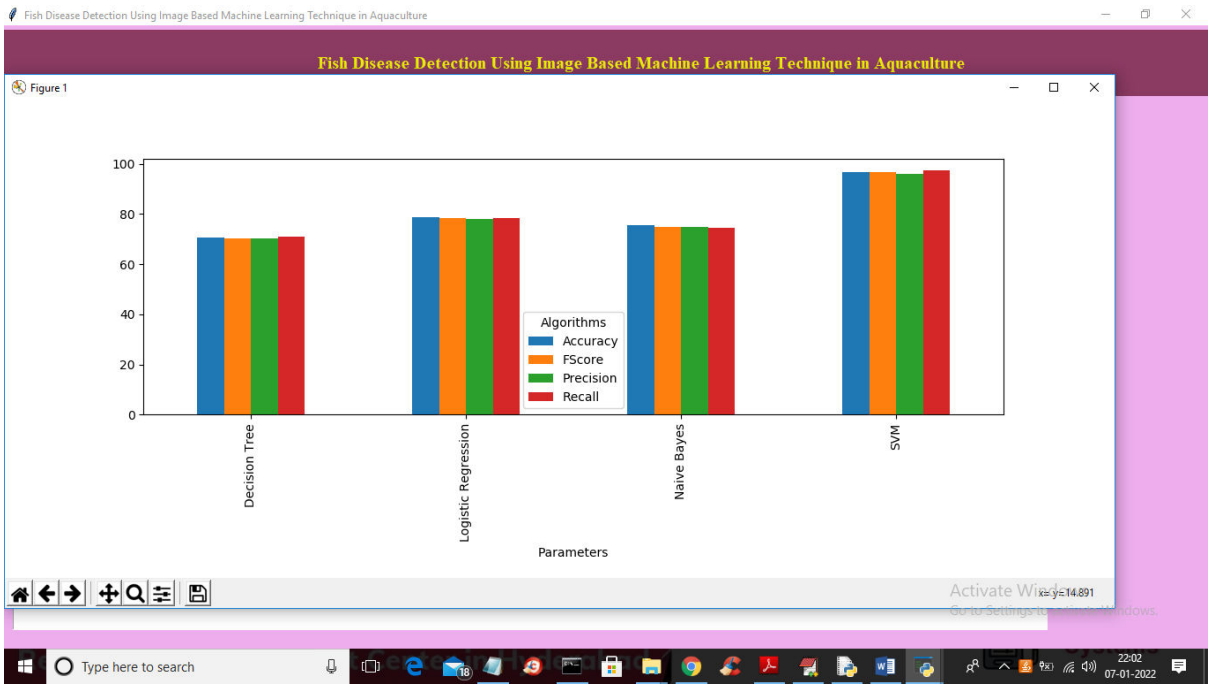
In above screen with logistic regression we got 78% accuracy and in confusion matrix we can see 28 predicted as Fresh fish image correctly and 6 are wrongly predicted. Now close above graph and then click on 'Run Naïve Bayes' button to train Naïve Bayes and to get below screen



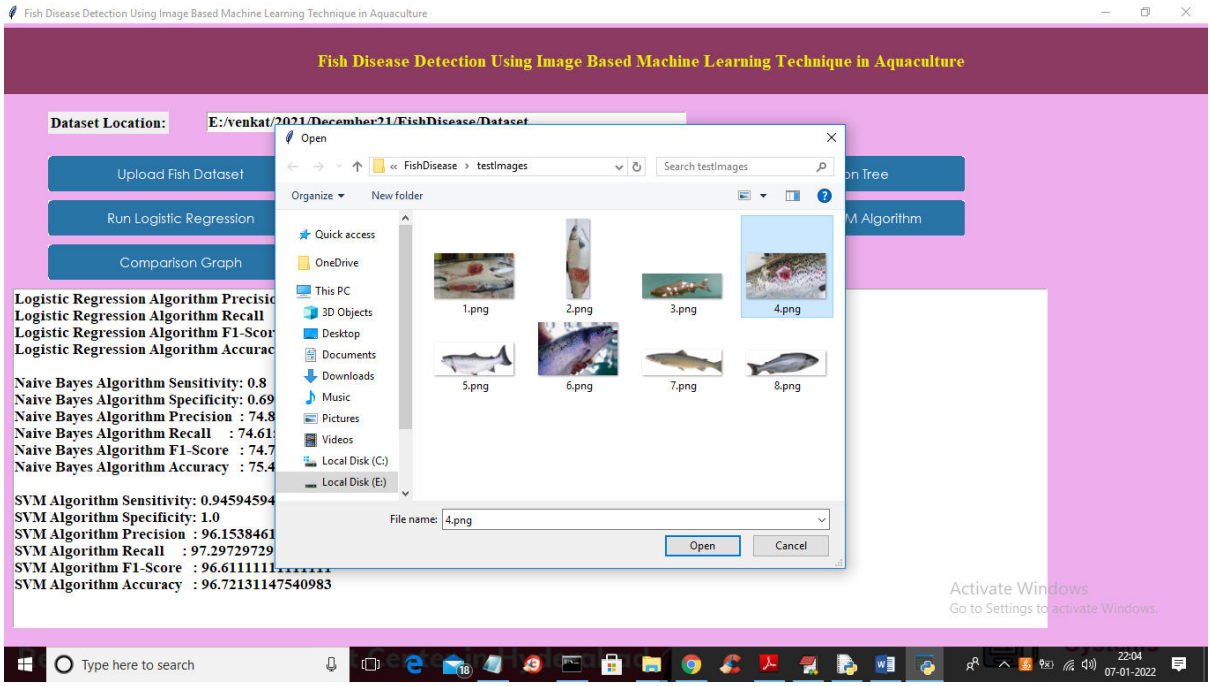
In above screen with Naïve Bayes we got 75% accuracy and in confusion matrix 28 images are correctly predicted ad fresh fish and 8 incorrectly predicted and now close above graph and then click on 'Run Propose SVM Algorithm' button to train SVM and to get below output



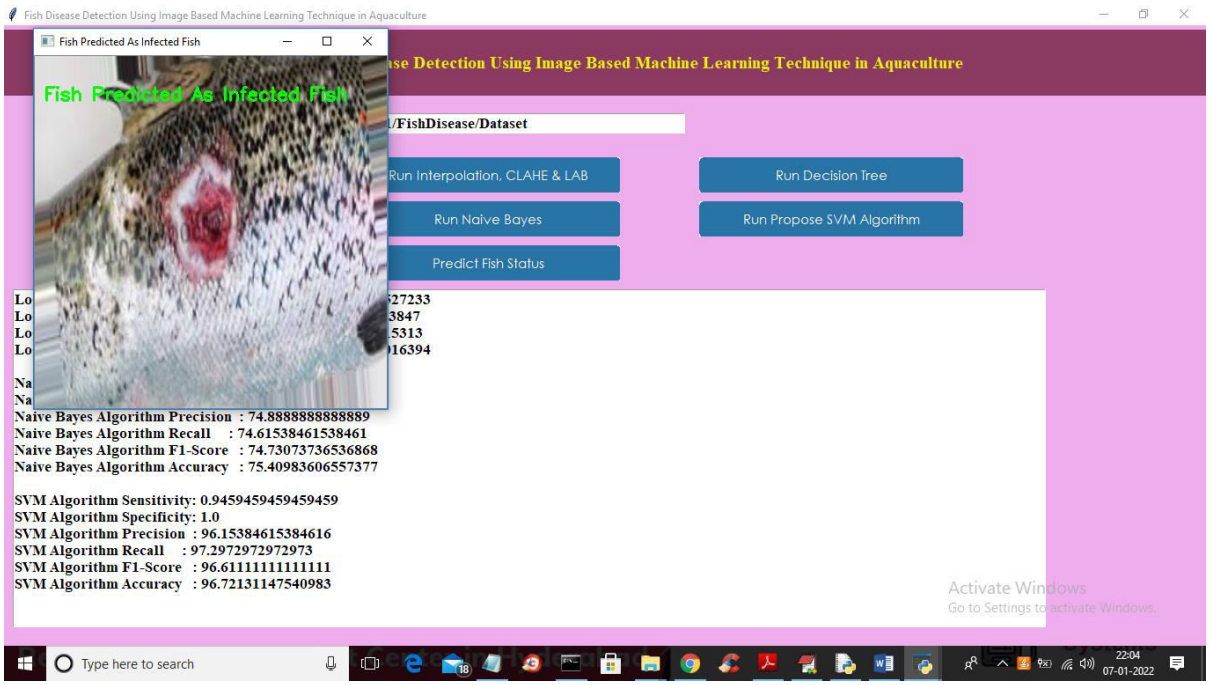
In above screen with SVM we got 96% accuracy and in confusion matrix graph we can see all 35 are correctly predicted as fresh fish and 0 incorrectly predicted. So SVM is better than other algorithms. Now click on 'Comparison Graph' button to get below comparison graph



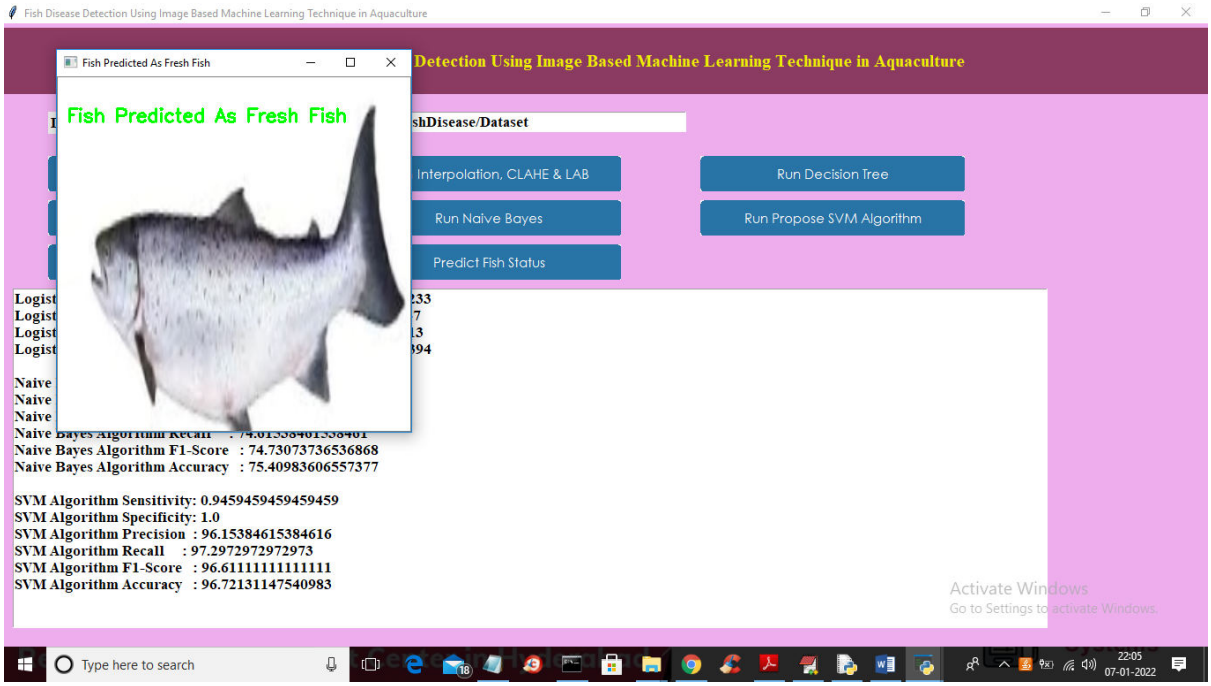
In above graph x-axis represents algorithm names and y-axis represents accuracy, precision and other metric in different colour bar. In above graph we can see SVM got high performance and now close above graph and then click on 'Predict Fish Status' button to upload test image and to get prediction



In above screen selecting and uploading '4.png' file and then click on 'Open' button to load image and to get below prediction



In above screen fish predicted as INFECTED FISH and now test other images



In above screen uploaded fish predicted as FRESH

Similarly you can upload other images and test

7. CONCLUSION AND FUTURE SCOPE

FOR IDENTIFICATION OF FISHES:

As a result of testing the model, we got a very accuracy of 50% of correct classification samples after 15 epochs for fish diseases and 80% of correct classification samples after

30 epochs for detecting types of fishes. The only drawback was less in accuracy because less data to train.

FOR IDENTIFICATION OF DISESES:

We got a accuracy of 90% by using —MOBILE NET v2‖ algorithm, the only drawback of this process is training the algorithm i.e., it takes almost 40 minutes of the time.

8. REFERENCES

[1] A. K. Sahoo,Shaveta Malik,‖Image Processing Techniques for Identification of Fish Disease‖,

[2] Y Wei-bo,Du Zun-sheng,‖An improved Kirsch human face image edge-detection method based on Canny algorithm‖,