

# A Novel Approach to Fish Disease Diagnostic System based on Machine Learning

<sup>1</sup>Shaveta Malik, <sup>2</sup>Tapas Kumar, <sup>3</sup>A.K Sahoo

<sup>1</sup>Research Scholar, School of Computer Science, Lingaya's University, Faridabad, Haryana

<sup>2</sup>Associate Dean, School of Computer Science, Lingaya's University, Faridabad, Haryana

<sup>3</sup>Scientist, Biodiversity Lab, Reverie Ecology and Fisheries Division, Barrackpore, Kolkata

Shavetamalik687@gmail.com, kumartapas534@gmail.com

## ABSTRACT

Real-Time identification automated system diagnoses fish disease i.e. Epizootic Ulcerative syndrome (EUS) which is caused by *Aphanomyces invadans*, a fungal pathogen. In this paper we propose a Real-Time fish disease diagnose system with better accuracy. In order to improve the accuracy we propose a combination (PCA-FAST-NN) which combine the Principle component analysis (PCA) with Features from Accelerated Segment Test (FAST) feature detector using Machine Learning Algorithm (Neural Network) i.e. (PCA-FAST-NN). The Experimentation has been done on the real images of Epizootic Ulcerative syndrome (EUS) infected fish database and implemented in MATLAB environment.

**Keywords:** Epizootic Ulcerative syndrome (EUS), Principle component analysis (PCA), Features from Accelerated Segment Test (FAST), Neural Network

## 1 Introduction

Fish Disease is one of the major deterrents in the aquaculture sector resulting in loss of more than 10 billion to the Indian agricultural export value. Routine diagnosis is an important part of the fish health management. Epizootic Ulcerative syndrome (EUS), a serious fish disease problem in India and in many countries including UK, Australia, Japan, Thailand and Pakistan. The outbreaks occur during early post-monsoon season especially during winter. It is noted that temperature is a critical factor for the fungal growth and fish mortality. The disease is clinically diagnosed as deep ulcers and heavy mortality caused by *Aphanomyces invadans*, a fungal pathogen. However, these ulcers are frequently misidentified as a causative agent of bacteria and virus. Due to the misidentification and high mortality rate of the fishes and to prevent the loss of agriculture, the software going to be develop will diagnose accurately the disease (EUS) and to prevent the loss of farmers and infact to prevent the loss of Indian economy. This paper proposed a combination of technique which gives better accuracy and compared with the existing combination. In proposed combination technique which combine the Principle component analysis (PCA) with Features from Accelerated Segment Test (FAST) feature detector and then classify through Machine Learning algorithm which is neural Network. It is a fast and accurate method to detect the disease in less time, late detection and diagnose leads to loss of fish as it is a food of millions of people in the world. Hitesh Chakravorty et al. [1] proposed a system in which the image of the diseased fish recognizes by using PCA method and diseased area segmentation of fish image based on color features with K-means clustering. HSV images and Morphological operation open for better accuracy to diseased area detection and measurement. It has taken only four Epizootic Ulcerative Syndrome (EUS) diseased fish images as a case study to evaluate the proposed approach

and in the experimental results clear indication of the effectiveness of proposed approach to improve the diseased identification with greater precision as well as correctly compute diseased area. In which features are extracted and processed by PCA to form the feature vector after that classify them according to the Euclidian distance. Valentin Lyubchenko et al. [2] In the paper choose or select the color markers to detect the infected area or to distinguish the normal and infected area. The main disadvantage of this methodology is the appearing of false point marked as infected areas due to the automatic allocation of the selected color which may affect the calculation. On the contrary, this methodology has many advantages such as: the ability to change the size of the marker when selecting the colors in the image segmentation to avoid any appearance of false points, the ability to analyze fish skin areas that are difficult to identify visually and the ability to isolate areas of the skin with varying levels of destruction. Jeyanthi Suresh et al.[3] In the paper, proposed a method which automatically recognizes the human activity from input video stream with Histogram of Oriented Gradient features (HOG) and Probabilistic Neural network (PNN) classifier and In which features are extracted from input video frames by HOG features detector and are concatenated to form an action pattern. The Probabilistic Neural network (PNN) is used as a classifier for classifying the actions of supplied video. Experiments are conducted on KTH database and gives better performance in terms of 100% recognition rate for training set and 89.8% accuracy for test set. SVM classifier is used for comparison and measuring performance of each feature set with respect to two classifier.

## 2 Methodology

In this approach, first the database is created by extracting image features from the real images of the fish and storing them in a data base. For any input fish image features will be found to match with the database. Training sample and testing sample has to be taken and the output will be identified images which is present in the database.

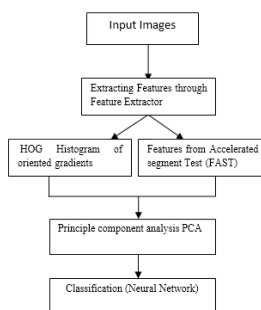


Figure1. Flow Chart of the System

In the above flow chart it explains the sequence which combines the feature detector with the classifier. In proposed system put the images as an input and then do the pre-processing, applied the morphological operations i.e. convert the image into grey scale, noise removal, edge detection and then apply the Feature extractor which extract the features and then PCA will apply which reduce the dimensions after the feature selection classifier (Neural Network) will apply which classify the images into EUS disease fish and Non-EUS disease..

### 2.1 Histogram of Oriented Gradients (Hog)

Dalal and Triggs et,al[12]HOG is feature descriptor and which is used for the object detection. It uses gradient orientation in localized portions of an image. The scanning of each image is based on detection window. In which the window is divided into cell and each cell accumulation a Histogram of Oriented Gradients over the pixels of each cell as shows in Figure 2 ,For better illumination, measure the local

histogram energy over block as a result all the cell normalize in block. It computes gradient orientation and magnitude and it also computed horizontal and vertical gradient. It is a technique which discovers shapes within the image and extract feature for all location in the image or region of interest (ROI).

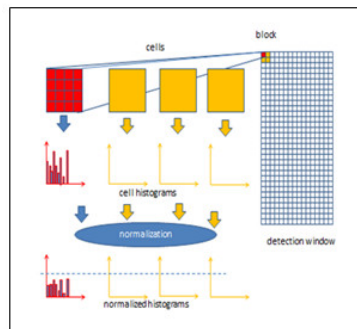


Figure2. HOG Algorithm Implementation scheme

## 2.2 EUS (Epizootic Ulcerative Syndrome) Infected Fish Image:-



Figure 3. a) Original Image of EUS Infected Fish



Figure 3 b) Applying HOG on EUS Infected Fish

HOG (Histogram of Oriented Gradients) applied on the EUS infected fish image as shown in the Figure 3 (b)

## 2.3 Principle component analysis (PCA)

PCA is a statistical method for finding correlations between features in data. It is linear technique for dimensionally reduction. It is used for reducing dimensionality of data by eliminating non-useful information from the dataset and it is frequently used in both image processing and machine learning technique. It chooses the best principle component for training data for each class or chooses first few components so that the dimensionality of the transformed data is reduced. It has been applied to the broad class of computer vision problems including feature selection object recognition etc. It is less sensitive to different training data sets. PCA (Principle Component Analysis) applied in the paper as it reduces the dimensionality of data after extracting the features through HOG (Histogram of oriented Gradients).

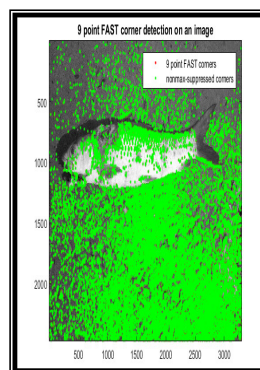
## 2.4 Features from Accelerated Segment Test (FAST)

It is a corner detection algorithm which is based on the intensity which minimizing the processing data and it is also used in image matching. Fast (Features from Accelerated Segment Test) relies on a corner

to response function and robustly detect corner in a given scheme. It is better than many other well-known feature extraction. It detects the interest point in the image. Interest point should be ideally repeatable between different images. The basic reasons behind to develop an interest point detector is using in real time frame rate application.



**Figure 4 a) Original Image of EUS Infected Fish**



**b) FAST Apply on EUS Infected Fish**

## 2.5 Neural Network

A neural network is based on the processing of information that has been developed as generalizations of mathematical model matching human cognition [10, 12]. The Neural Network is based on the belief and working on human brain by making or attaching the right connections, can be imitated using silicon and wires as living neurons and dendrites. The human brain consists of 100 billion nerve cells called neurons. They are connected to other thousand cells or neurons by Axons. Stimuli from external environment or inputs from sensory organs are accepted by dendrites. These inputs create electric impulses, which quickly travel through the neural network. A neuron can send the message to other neuron to handle the issue or does not send it forward.

### 2.5.1 Machine Learning in ANNs

ANNs are powerful tool and capable of learning and they need to be trained. There are some learning strategies –

- Supervised Learning –It is a type of Machine Learning algorithm and it uses a training dataset. It involves supervision or like a teacher which gives supervision. For example, pattern recognizing. It is based on the labeled training set and the class of each piece of data is known. Class label is pre-determined which will be provide in the training phase.
- Unsupervised Learning – It is also a type of machine learning algorithm and it is not dependent on the supervision, in which there is no class label and not having pre-determined class lables.For example, searching for a hidden pattern. Clustering i.e. where grouping will be done and it is based on the similarity and data have no target attribute.
- Reinforcement Learning – It is based on the input data, Reinforment Learning agent learn by interacting with its environment .It relies on being able to monitor the response of the action taken and action will be measure against a definition of a reward .

### 3 Proposed Methodology

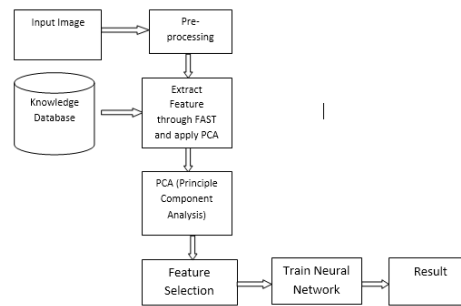


Figure: 5 Proposed Methodology

In the above proposed system as shows in Figure 5 ,Acquire the image or Input the Image those collected form NBFGR(National Bureau of Fish Genetic Resources),Lucknow then apply the morphological operations(convert the image into gray, removal of noise, segmentation) After that apply the feature extractor which is FAST which detect the interest point and interest point is ideally repeatable between different images, once the feature has been extracted from FAST it will be reduce by PCA which is used to reduce the dimensionality after selection of the features apply the classifier which is Neural Network and train it for detecting the fish disease as a result it gives better accuracy by matching the Training dataset with the testing dataset and get the accuracy. The proposed method automate the process and reduce the time of diagnose.

➤ **Steps in Proposed Methodology:**

- a) **Image Acquisition`**: - It is the process in which acquired or collect the image and converted to the desired output format according to the system.
- b) **Image Pre-Processing**:- The pre-processing steps are performed on the acquired image. Increase the contrast of the image , convert the image into gray scale, removal of noise and segmentation.
- c) **Feature Extraction and Recognition**:-In which to detect the symptom of the disease. Once the captured image is pre-processed. Extract the feature from the feature extraction algorithm which is FAST (Features from Accelerated segment Test) for recognition. To diagnose EUS, Infected images are loaded into the database and features are extracted after that will be treated by PCA (Principle Component Analysis) to reduce the dimensionality for the better accuracy.
- d) **Classification**: -After the feature selection classification will do by classifier which classifies the disease fish and non-disease fish. The dataset to be divided into training validation set and testing validation set which will be feed into the Neural Network. The Proposed Methodology to detect or diagnose whether the disease is present or absent in it. It is a Fast and accurate method for detecting and classification of the fish disease images and it is efficient and easy way to detect the EUS disease. After applying the PCA .The experiment results indicate that the proposed combination is fast and efficient in recognition and it gives better accuracy.

## 4 Experimentation & Results

### 4.1 Fast-PCA with Neural Network

**Accuracy:** Calculate the portion of True positive and true negative in all evaluated cases. It is its ability to differentiate the infected fish and healthy fish.

$$p = \frac{TP}{TP + FP}, \quad r = \frac{TP}{TP + FN}. \quad (1)$$

- Precision  $p$  is the number of correctly classified positive examples and divided by the total number of examples that are classified as positive.
- Recall  $r$  is the number of correctly classified positive examples and divided by the total number of actual positive examples in the test set.

#### 4.2 Confusion Matrix of FAST-PCA-NN

As shows in figure 6, the confusion Matrix shows the 86.0% accuracy through the classifier.

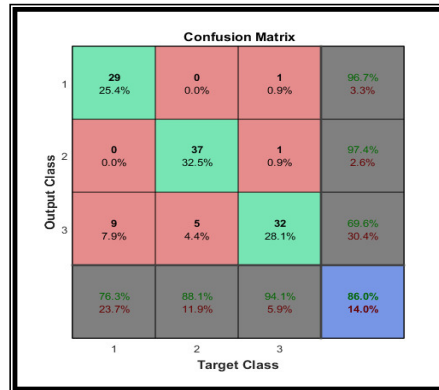


Figure: 6 Confusion Matrix of (FAST-PCA-NN)

The Experimentation which has been done on the MATLAB environment to detect or diagnose the EUS infected and Non-EUS infected disease fish. The proposed combination gives better results and accuracy as compared to existing.

#### 4.3 Cross Entropy Vs Epochs (FAST-PCA-NN)

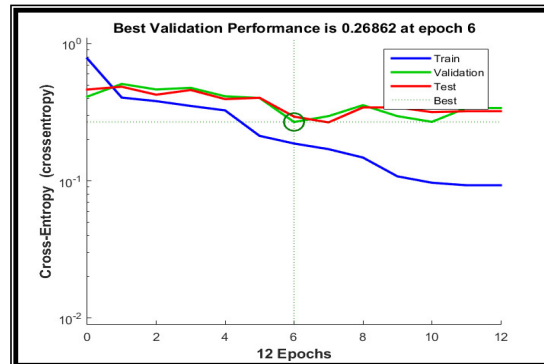


Figure: 7 Cross-Entropy Vs epochs

The above figure shows the cross entropy Vs epochs and gives the best validation performance which is 0.26863 at epoch 6.

#### 4.4 Confusion Matrix of HOG-PCA-NN

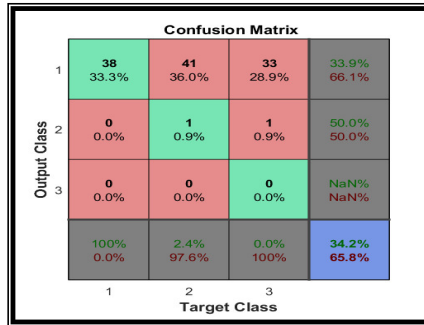


Figure: 8 Confusion Matrix of (HOG-PCA-NN)

The above figure 7 shows a confusion matrix of HOG-PCA-NN which gives 65.8% accuracy. It is not better than the proposed combination.

#### 4.5 Cross Entropy Vs Epochs (HOG-PCA-NN)

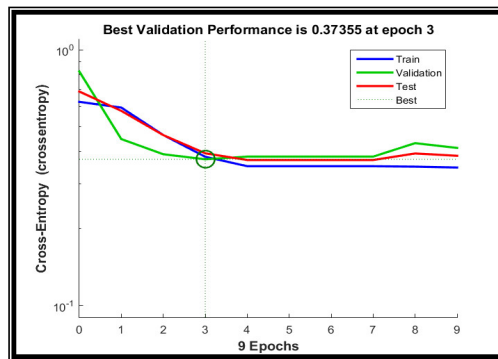


Figure:-9 Cross Entropy Vs Epochs

As shown in figure 9 which is cross entropy Vs epochs of HOG-PCA-NN, it gives the best validation performance at epoch 3 which is 0.37355.

#### 4.6 Comparison between the accuracy of HOG-PCA-NN and Proposed FAST-PCA-NN

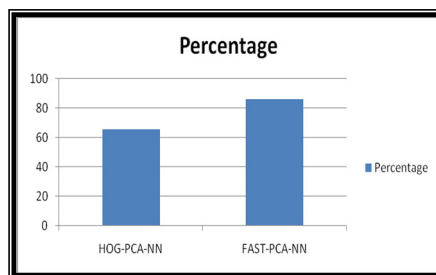


Figure:-10 Accuracy between HOG-PCA-NN and FAST-PCA-NN

As shows in Figure 10, the proposed combination gives better accuracy as compared to the previous which is (HOG-PCA-NN). The showing the comparison between the classification accuracy of HOG-PCA-NN and proposed combination which is FAST-PCA-NN.

## 5 Conclusions

The proposed combination of FAST-PCA-NN gives better classification accuracy than the existing combined technique which is PCAs-HOG-NN to recognize the EUS infected disease fish as it is misidentified to detect. It is fast and efficient method to recognize and detect the EUS disease as compared to the Traditional Method. The Matlab simulator results based on the real images of EUS infected fish disease dataset. The proposed combination (FAST-PCA-NN) gives 86.0% accuracy.

## ACKNOWLEDGEMENT

The EUS infected images have been collected from National Bureau of Fish Genetic Resources (NBFGR, Lucknow) and ICAR-Central Inland Fisheries Research Institute (CIFRI), Kolkata. Thanks to Dr. A.K Sahoo(CIFRI,Kolkata) and Dr .P.K Pradhan(NBFGR,Lucknow).

## REFERENCES

- [1] Hitesh Chakravorty, Rituraj Paul & Prodipto Das , *Image Processing Technique To Detect Fish Disease*, International Journal of Computer Science & Security (IJCSS), Volume (9) : Issue (2) : 2015 121
- [2] Jeong-Seon Park, Myung-Joo Oh, and Soonhee Han, *Fish Disease Diagnosis System Based on Image Processing of Pathogens Microscopic Images*, Frontiers in the Convergence of Bioscience and Information Technologies 2007
- [3] Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., Ford, S. E. (2014). Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., Ford, S. E. (2014). Climate change influences on marine infectious diseases: implications for management and society. Annual review of marine science, 6, 249-277.
- [4] Lafferty, K. D., Harvell, C. D., Conrad, J. M., Friedman, C. S., Kent, M. L., Kuris, A. M., Saksida, S. M. (2015). *Infectious diseases affect marine fisheries and aquaculture economics*. Annual review of marine science, 7, 471-496.
- [5] Afferty, K. D., Harvell, C. D., Conrad, J. M., Friedman, C. S., Kent, M. L., Kuris, A. M., Saksida, S. M. (2015). *Infectious diseases affect marine fisheries and aquaculture economics*.,Annual review of marine science, 7, 471-496.
- [6] Valentin Lyubchenko,Rami Matarneh,Oleg Kobylin,Vyacheslav Lyashenko, *Digital Image Processing Techniques for Detection and Diagnosis of Fish Diseases*, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 7, July 2016
- [7] Narasimha-Iyer, H., et al., *Automatic Identification of Retinal Arteries and Veins From Dual-Wavelength Images Using Structural and Functional Features*. Biomedical Engineering, IEEE Transactions on, 2007. 54(8): p. 1427-1435.
- [8] Daoliang Lia, Zetian Fua, Yanqing Duanb,*Fish-Expert: a web-based expert system for fish disease diagnosis*, Expert Systems with Applications 23 (2002) 311–320.
- [9] Nayak, K. K., Pradhan, J., Das, B. K. (2014). *Original Research Characterization, pathogenicity, antibiotic sensitivity and immune response of Flavobacterium columnare isolated from Cirrhinus mrigala and Carassius auratus*. Int. J. Curr. Microbiol. App. Sci, 3(11), 273-287.



- [10] J.-Y.Chang and J.-L. Chen, "*Automated facial expression recognition system using neural networks*", Journal of the Chinese Institute of Engineers, vol. 24, no. 3, (2001), pp. 345-356.
- [11] M. Rizon, M. F.Hashim, P. Saad, S. Yaacob, "*Face Recognition using Eigenfaces and Neural Networks*", American Journal of Applied Sciences, vol. 2, no. 6, (2006), pp. 1872-1875.
- [12] Navneet Dalal, Bill Triggs, "*Histograms of Oriented Gradients for Human Detection*"International Conference on Computer Vision & Pattern Recognition (CVPR '05), Jun 2005, San Diego, United States.