A Novel Feature Extractor Based on the Modified Approach of Histogram of Oriented Gradient

Shaveta Malik[1(✉)], Archana Mire[1(✉)], Amit Kumar Tyagi[2(✉)], and Vasudha Arora[3(✉)]

[1] Terna Engineering College, University of Mumbai, Mumbai, India
shavetamalik687@gmail.com, archanamire@ternaengg.ac.in
[2] School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai 600 127, Tamil Nadu, India
amitkrtyagi025@gmail.com
[3] G.D Goenka, Gurugram, India
Vasudharora6@gmail.com

Abstract. In image processing, the goal of feature extraction is to extract a set of effective features from the raw data. Feature extraction starts from an initial set of measured data and builds derived values i.e. features intended to be informative and Non-redundant. The paper is based on the novel feature extraction approach for the detection of Epizootic Ulcerative Syndrome (EUS) fish disease which is misidentified among people. The EHOG (Enhanced Histogram of Oriented Gradient) which is a proposed feature Extractor to extract the features or information. The paper discuss its comparison with other existing techniques with different parameters. The Evaluation results shows that the EHOG is better in every parameters and also gives better accuracy and efficiency of the model which recognizes the disease.

Keywords: Epizootic ulcerative syndrome (EUS) · Histogram of oriented gradient (HOG) · Neural network · Fish disease

1 Introduction

The seasonal epizootic condition of great importance in wild and farmed fresh and brackish water fish is Epizootic Ulcerative Syndrome (EUS). It is considered as an indistinguishable now from red spot disease and it is observed first in eastern Australia [1]. The fungus *Aphanomyces* invader (A.piscicida in Japan) is the causative agent of EUS, when it invades the body, it causes intense liquefactive necrosis of muscle tissue in some cases the hyphae extend into the visceral organs. Often mass mortality is observed as an initial sign of distinct dermal lesions, including Ulcers. Fishes surviving from the diseases may get burn like marks, red-spots, blackish and deeper ulcers with red centres and white and it is usually have lesions of varying degree of severity. Especially in Asia-Pacific region and Africa due to spread of the disease has led to substantial damage to the livelihood of the fish farmers and fish resources [2]. It is hard to differentiate among fishes having EUS disease and ulcers. Red spot is not sufficient

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12254, pp. 753–770, 2020.
https://doi.org/10.1007/978-3-030-58817-5_54

to differentiate [3–5]. Nevertheless by using image processing techniques and pattern recognition, it will be easy to recognize the EUS disease pattern of fish. Correct recognition will prevent the mortality of the fish. This paper proposed a new (modified) approach of feature extractor that is Enhanced Histogram of Oriented Gradient (EHOG) which extract the multiple features and provide more information and better accuracy by using the classification algorithm. The method of detecting the disease is divided into three steps i.e. pre-processing, resizing, normalization and then feature extraction to get the features after that classification will be done. After extracting the features it will be easy to classify the images. However The EHOG which is an enhanced feature extractor and it gives more enhanced and better results in every parameter.

1.1 Different Type of Features

The main task of feature extraction is to extract the relevant information. When the input data is too large and it takes too much time to process then it must have redundancy. It transform into reduced features that is named as feature vector. There are number of features as given below:-

Shape Features: For image content description shape is an important feature. The description of the size content cannot be defined exactly because it is difficult to measure the similarity between shapes. It has two parts one is contour based and another is region based. Here, Contour based is used for the boundary and region base is used for the whole image. As the dimensionality increases, the amount of training data increases. Moreover, there is an impetus to combine the features and to produce a feature vector as about shape & size and correlation between different features.

Edge and Boundary Features: It is a fundamental (basic) problem in image processing. In image processing, edges in the images have strong intensity and the quality of image is affected by the variations in intensity but the information will be preserved. The basic properties such as area, perimeter and shape can be measured easily only if the edges are identified accurately. In segmentation also the estimation of boundary edges are used.

Texture Feature: It is very significant (essential feature) property of image which helps in the process of retrieval. It maintains the information about the variation of the intensity, with the regular interval. The texture has the structure of the arrangement or repeated pattern of information. It is divided into two methods one is spatial and other is spectral texture. There are some pros and cons of it. In spatial it is easy to understand and can be extracted the information from any shape with losing the information but it is sensitive to noise. In spectral texture has less computational power and it is robust in nature but it need region of square image that can be of sufficient size.

The organisation of the paper is follows as: In Sect. 2, tells about the background and literature review of the related paper. Further, Sect. 3 discusses about the motivation why instant or correct detection of EUS disease is required today. Section 4, discusses about the basic algorithm of histogram of oriented gradient. Later, Sect. 5 discusses about the methodology which tells about general method of the process. Further, Sect. 6 discusses about the enhanced histogram of oriented gradient which is

more efficient as compared to the basic Histogram of oriented gradient. In Sect. 7, discussion done about the classification with Neural Network algorithm and dataset. In Sect. 8 discussed about the analysis of our results. In last, this work is concluded with some interesting future scope.

2 Literature Review

The paper is based on the recognition [6] of the fish and it is a challenging task, it is based on the recognition of image using the coefficient correlation and then applied the HOG feature extractor then classified the fish images by using the SVM (Support Vector Machine). It is to detect the species of the fish on the boat of fishing. In which images were captured from boat cameras through various angles and that is to be used to prevent the endangered of fishing to fishes. The accuracy was 94% through the feature extractor which is HOG (Histogram of Oriented Gradient). The paper is based on the automated system [7] for the recognition of the images of color logo. In which color feature was used for recognition, for example, color moments and the feature extractor which had used Histogram of Oriented Gradients (HOG). In this, Classification of the logo images was done by using Support Vector Machine Classifier (SVM), i.e., the classifier used to recognize the logo image by Support Vector Machine (SVM). In which compared the HOG-SVM with other existing methods then the HOG-SVM approach was the fast in execution and easy in implementation and it gave 88.50% accuracy. In [8] authors focused the recognition of hand gesture using the different feature extraction techniques and classifier used is SVM. The proposed model was the hybrid of SIFT and HOG and classified through the SVM. In the age of computer the gesture of hand recognition is very important domain of the application of computer. The proposed model gave 97% of accuracy with the 10 gestures.

Further, in [9] authors proposed a method which automatically recognize the activity of human from the video stream using 'Histogram of Oriented Gradient (HOG) and classifier which is Probabilistic Neural network (PNN). The features of actions were extracted from input video frames. PCA was used for reduction of dimensionality. Experimentation was conducted on the KTH database and gave 89.8% accuracy for test set. The experimentation was implemented in MATLAB version 8.1.604 R2013a. The paper presented a novel method using PNN for action recognition and HOG features extractor [10] with PCA which for dimension reduction as the tool of feature extraction. This is for the applications of action recognition in fields like surveillance entertainment and healthcare systems. Further, in [11] authors presents a comparative study of the algorithm of matching, it define the efficiency level. In which FAST (Feature Accelerated Segment Test) algorithm is used and detects limited features and to the features was detected among the best one and detected features were with high contrast to the feature surroundings. The FAST algorithm works well with planar images and the fast too and the result came from the FAST in the matching efficiency 58.06%. Note that there are various attempts have been made in the previous decade with respect to this (discussed) problem, but none of approaches work efficiently. Further, some other mechanisms like Human Detection using Oriented

Gradients have been proposed [11, 12], also some other attempts have been made by many authors (in the past decade) in [10, 13–15].

In paper [16], author done an experimentation by ANN & SVM for the early detection of the lung cancer. In which ANN is ensemble with HOG for the prediction of lung cancer. Moreover it proposed a framework which has multiple biomarkers for the lung cancer and extracted the features from the Nucleosides sequences. HOG & LBP (Linear binary pattern) used to extract the features. ANN ensemble with ANN and gave 95.90% accuracy.

In paper [17] the author described the co-occurrence Histogram of oriented gradient (CO-HOG) that has been proposed to recognize the text in the scenes and compared with HOG (Histogram of oriented gradient). The weighted voting scheme has been used for character recognition in scene. The accuracy of proposed method was 80.6% and HOG with SVM classifier was 94.890%.

In summary, this section discuss about the performance or accuracy from different methods or techniques with different feature extractor and model. Now, next section will deal with the motivation why it is necessary to correct detection of EUS disease and what is the cause of increasing the mortality of the EUS fish disease.

3 Motivation

Fish is the livelihood of millions of people. Epizootic Ulcerative Syndrome (EUS) is one of the most serious aquatic diseases [18]. High mortality and fish rejection cause high losses to fish farmers and fishermen and the next concern is about the health due to the presence of ugly lesions and also abridged productivity of all susceptible fish species [19, 20]. The fish farmer's faced different restrictions when fish disease outbreaks occur in their farms.

There are number of factors which are the cause of fish mortality or outbreak, few reasons are listed here as:

- Lack of Knowledge on Treatment.
- Lack of Knowledge of fish disease.
- Lack of advisory services from government and non-government organization.
- Lack of Training Facility about fish disease treatment.

The above are the general factors for the mortality of fish.

4 Histogram of Oriented Gradient (HOG)

The Histogram of Oriented Gradients (HOG) is a feature extraction method which is used for the image classification [21], it computes the local gradient and magnitude with the overlapping of blocks. It focuses on the shape and size. It is for the application of the detection of person and it is focussed on where the feature in the image is quite large. Many methods or approaches believe on the Gaussian filtering methods but the HOG does not rely on the filtering Methods [22, 23].

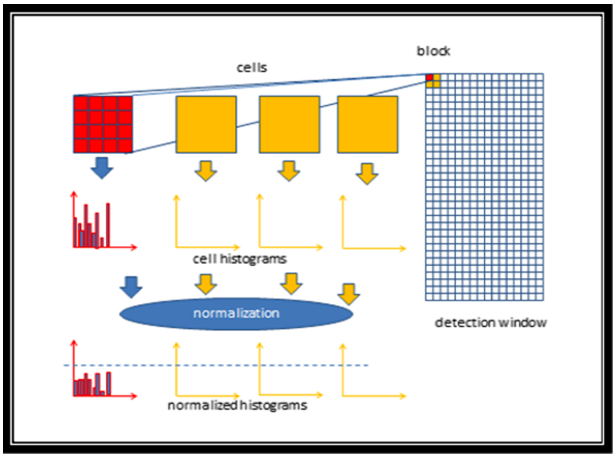


Fig. 1. Histogram feature extraction [16]

The Fig. 1 shows the image is divided into cells and then combine into block. The histogram shown in every block.

The algorithm works as follows:-

- Calculate the gradient of X and Y variable of each pixel of image.
- Divide the cell and after that combine it into the block.
- Calculate the gradient magnitude and direction with respect to the angle of each pixel of the image.
- Each cell is of specific dimension or cell size as it depending upon the gradient angle, allocate the gradient magnitude in predefined bin with the range of 20° only if the number of the bins will be 9.
- It should be as the range of 0 to 20, 20 to 40, 40 to 60, etc., and for those angles which are not in the centre of the bin, for example, 10, 30, 50, 70, etc. in the traditional approach HOG approach if the gradient with the angle 15° which is closer to the bin 1 and then 75% of its gradient will allocated to the bin 1 and 25% to the bin 2 [24, 25].
- In every cell of histogram of oriented gradient obtained with number of bins and the magnitude of every bin by calculate the magnitude of every bin by calculating the corresponding pixel of gradient and adding the interpolated gradient magnitude of all corresponding pixel.
- The cell can be group to form a block and the magnitudes of all histogram are normalized within the block.

Hence this section discuss about the basic algorithm of Histogram of oriented gradient, next section will discuss about the methodology about the general concept of classification.

5 Methodology

In methodology, the general concept of the classification is to extract the features from the feature Extractor after the pre-processing. In pre-processing, the number of the steps to be cover, i.e., binirization [26], Normalization etc. After that select the feature extractor which extract the features as useful information although the quality of the information depends on the feature extractor and then classify the EUS and Non-EUS through classification algorithm.



Fig. 2. Block diagram of methodology

The Fig. 2 shows the steps of methodology

Step 1: Extraction: The features from the feature extractor which extracts the features to extract the information after the pre-processing.
Step 2: Selection: After extracting the features then select the selected features from the feature extractor.
Step 3: Classification: Each object is represented by a feature vector for training and testing. In the problem of image classification, this feature vector is usually obtained using pixel-based method [27]. In the paper, have applied the feature extraction HOG from vector by applying the proposed feature extractor a feature vector V1 is obtained, then applied PCA (Principal Component Analysis) to the feature vector and form a new vector which is V2 as it has the lower dimensionality then apply the classification algorithm which is Neural Network.

Hence, this section discusses methodology. Next section discusses about modified feature extractor, i.e., Enhanced Histogram of Oriented Gradient (EHOG) in detail with a flow chart.

6 Proposed Feature Extractor (EHOG)

In the proposed methodology a modified approach which is Enhanced Histogram of Oriented Gradient as a new efficient feature extractor. It is expected to have better accuracy performance for EUS detection. It helps to enhance the features and improve the accuracy as compared to the other algorithms.

The flow chart of Modified Feature Extractor explain below:-

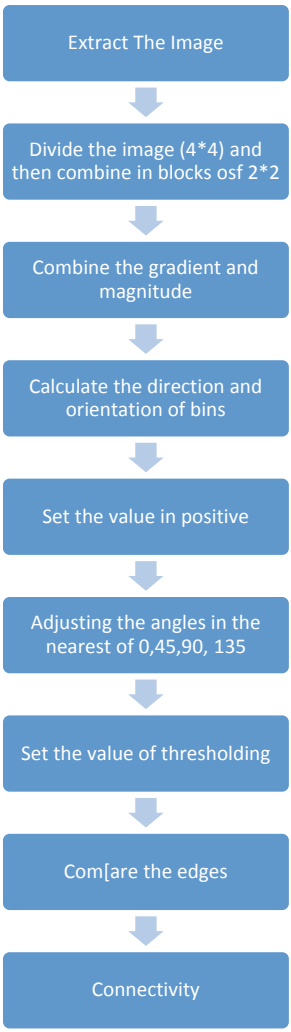
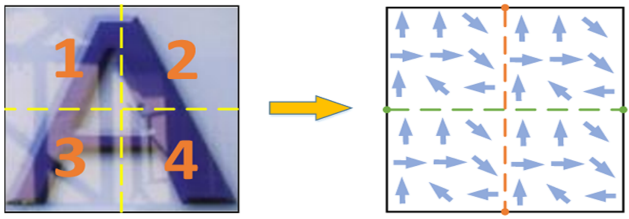


Fig. 3. Flow chart for enhanced histogram of gradient

The Fig. 3 explains the method to extract the features of EUS disease image through EHOG (Enhanced Histogram of Gradient).

Step 1: Divide the image into 4*4 for processing and then combine into 2*2 block with overlapping of blocks because in the paper taking the block size 2*2 and cell size 4*4 and it can be vary. Calculate the Block per image as showed in figure below.



(a) Sample image (b) Gradient Orientation

Fig. 4. a) Image divide into four parts [17] b) orientation of gradient [17]

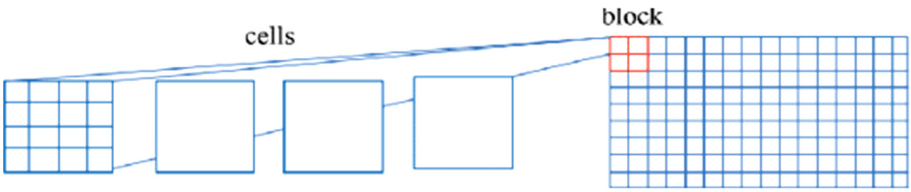


Fig. 5. Cell divided into blocks

In Fig. 4 and 5 Image is divided into four parts and cell is divided into blocks.

Calculation to find block per image: - Block per Image = Absolute value ((size (I)/cell size) Block size)/(Block size-block overlap) + 1).

The Fig. 3 explains the method to extract the features of EUS disease image through EHOG (Enhanced Histogram of Gradient)

e.g. B = Abs ((50*50)/4*4)2*2)/(2*2 − 1*1) + 1) (In case size of image is 50*50, then there will be 192 blocks per image) where I = Image.

Step 2:- After that calculate the magnitude and gradient while calculating the gradient in terms of X and Y direction, gradient will vary according to the values.
Step 3:- Calculate the orientation and direction of bins while calculating the theta. (Orientation is divided into bins).

Note:- If alpha $\geq$ 0 then Value is equal to Alpha if Alpha < 0 then Alpha = Alpha + 360, i.e., Alpha < 0 then b = 360 + Alpha where alpha is for angle then adjusting the angle into 0, 45, 90, 135°

Step 4:- Set the value of thresholding if the value less than T_{Low} then it is set to 0 and if the value greater than T_{high} then it is set to 1, all the values will be in the 0 and 1 after converting it into 0 and 1 according to the range then connect and compare the edges or corners.

6.1 Algorithm of Enhanced Histogram of Oriented Gradient (EHOG)

- Extract the EUS image.
- Divide the image into cells of size 4*4 and then combine into blocks 2*2

- After that compute the gradient and find the magnitude

$$\text{Gradient Magnitude} = \sqrt{g_x^2 + g_y^2} \tag{1}$$

$$\text{Gradient Direction} - \theta = tan^{-1}\left[\frac{g_y}{g_x}\right] \tag{2}$$

Derivate with respect to gradient in the X direction - Derivative with respect to gradient in the Y direction (refer Eq. 1 and 2)

- Calculate the directions and orientations of bins.
- For [− 180,180] to [0,360] if alpha $\geq$ 0 then Value is equal to Alpha if Alpha < 0 then Alpha + 360, i.e., Alpha < 0 then b = 360 + Alpha
- Adjusting direction nearest to 0, 45, 95 and 135°, i.e., if the angle is nearer to 45 then it comes in the 45°. **//All the value should set in positive.**
- Set the value of thresholding, i.e., T_{Low}- 0.075, T_{high}- 0.175.
- Calculate the 8-Connected Blocks.
- Calculate the edges or corner

Hence this section discuss about the Modified feature extractor which is Enhanced Histogram of Oriented Gradient (EHOG). Next section will discuss about the Classification and Confusion Matrix.

7 Classification

For classification divide the dataset into two parts one is training set and another is testing data set then apply the neural network algorithm which classify the dataset into EUS fish image and Non-EUS fish image.

Neural Network

Neural Network is organized in layers all layers are connected with each other with neurons [28]. Input will be given to the input layer or patterns are recognize through the input layer but the actual processing is done through hidden layer with the weights and hidden layer is connected with output and output layer gives output.

Figure 6 shows that layer 1 is the input neurons, all the neurons are interconnected with hidden layer and then goes to the output.

Different feature extractor have been applied with Neural Network algorithm in the paper and calculate the values or analyse the value after classification.

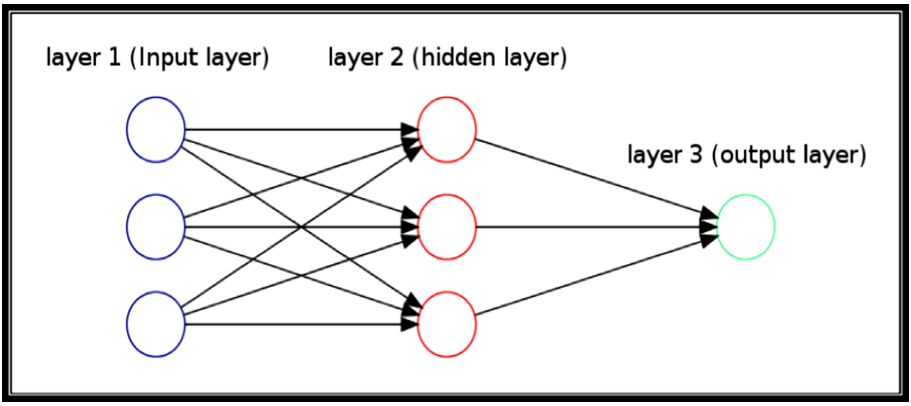


Fig. 6. Neural network

7.1 Sample of Training and Testing Dataset

The sample of "EUS (Epizootic Ulcerative Syndrome)" infected fish used in experimentation are the real images and collected from "(NBFGR, Lucknow) and ICAR-Central Inland Fisheries Research Institute (CIFRI), Kolkata".

Figure 7 shows the images of EUS (Epizotical Ulcerative syndrome) infected disease.



Fig. 7. Sample of training dataset (EUS Infected Fish) [16]

Classify the EUS and Non-EUS infected fishes through Neural network algorithm. The dataset is divided into training set which is Ttraining and testing set which is Ttesting. Portioning the data set into 70% training, 15% validation, and 15% testing. The data training set was used to train the network while the validation set was used to measure the error and the network training stops, when the error starts to increase for

the validation dataset. Furthermore, to get the better results, train the neural network many times and get the average of classification accuracy. In which 10 cross validation has been used.

7.2 Confusion Matrix

Confusion Matrix gives the prediction results on classification. In data set, Total number of 80 images taken for training and testing of EUS and Non- EUS images. The images have been collected from CIFRI-Kolkata (research institute of fish) and NBFGR, Lucknow (research institute of fish disease). While classification the confusion matrix has created to find out the positive and negative results with respect to the actual and predicted class.

In Fig. 8 shows about the general confusion Matrix which help in to calculate number of parameter with respect to positive and negative values of True and False.

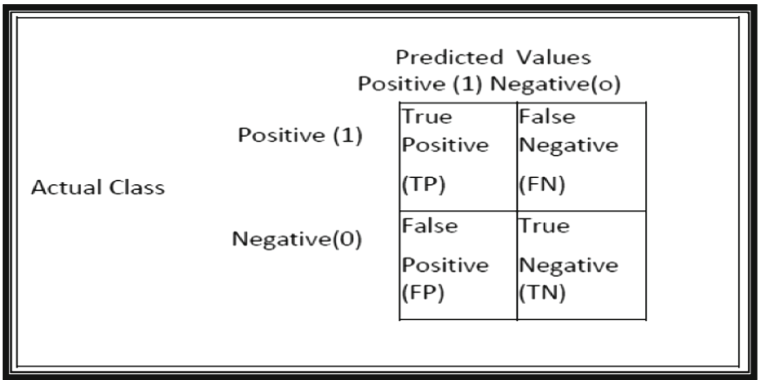


Fig. 8. Confusion matrix [27]

7.3 Confusion Matrix of EHOG with Neural Network Algorithm

Neural Network is used to classify the data and Enhanced Histogram of oriented gradient feature extraction has been used with PCA (principal of component analysis). It is used for the dimensionally reduction.

In Fig. 9 it is shown that how many images are correctly classified and how many are not correctly classified. It shows the accuracy also which is 98.8%. The first of the two diagonal cells which is in green color shows the number of correct classification by the trained model e.g. 39 fish images are correctly classified as EUS images that corresponds to 48.8% of total 80 images of fish. Similarly 40 fishes are correctly classified as Non-EUS image in diagonal opposite cell (green color in row 2). This corresponds to 50.0% of fishes. In fish column red colour cell one fish image which is Non-EUS, incorrectly classified as EUS and corresponds to 1.3% from total number of fishes images. In red color cell second column, zero number of the EUS images classified incorrectly as Non-EUS and this corresponds to 0.0% from total number of

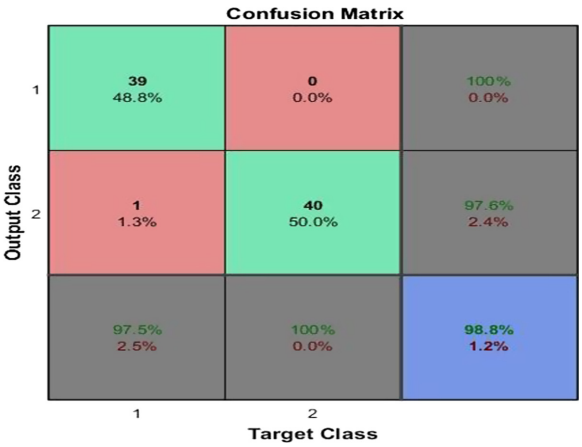


Fig. 9. Confusion matrix of EHOG-PCA-NN (Color figure online)

fish images. Out of 40 fish images in the first column EUS fish images prediction is 97.5% is correct and 2.5% is wrong. In second column last row out of 40 number of images. Non -EUS image prediction is 100% correct and 0.0% is wrong. In first row out of 39 EUS fish images 100% are correctly predicted as EUS and 0.0% is predicted as Non-EUS. In second row out of 41 Non-EUS images, 97.6% is correctly classified as Non-EUS and 2.4% is classified as EUS.

8 Performance Analysis

The Proposed feature extractor which is Enhanced Histogram of Oriented Gradient compared with the other feature extractor in terms of different parameters [29, 30] and it finds that the EHOG is better in every parameter as compared to the others according to analysis shows in the paper.

8.1 Analysis Performance Through Different Parameter

There are different parameters have been taken to compare the different feature extractor:

1) Accuracy
2) Precession
3) Recall
4) Specificity.

- **Calculation of Every Parameter**

1) **Accuracy:-** The accuracy rate can be calculated by the formula.

$$\text{Accuracy}: \quad \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{FN}+\text{TN}} \tag{3}$$

Where TP = True Positive, TN = True Negative, FN = False Negative and FP = False Positive.

a) Positive (P): It observed that it is positive (e.g., it is a fish)
b) Negative (N): It observed that it is not positive (not a fish)
c) True Positive (TP): It is observed that it is positive, and it is predicted to be positive.
d) False Negative (FN): It is observed that it is positive, but it is predicted as negative.
e) True Negative (TN): It is observed that it is negative and it is predicted as negative.
f) False Positive (FP): It is observed that it is negative and it is predicted as positive.

e.g. Taking the example of confusion matrix of EHOG-PCA-NN (Fig. 9) to find the accuracy refer formula no 3:

$$\text{Accuracy:-} \quad \frac{30+40}{80} = 98.8\% \tag{4}$$

2) **Precision**:- It is defined as the total number of correctly classified example of positives is divided by total number of positive predicted.

$$\text{Precision:-} \quad \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{5}$$

e.g. Taking the example of confusion matrix of EHOG-PCA-NN (Fig. 9) to find the Precision refer formula no 5:

$$\text{Precision:-} \quad \frac{39}{39+1} = 97.5\% \tag{6}$$

3) **Recall:-** It is defined as the correctly positive classified samples or examples is divided by True positive samples/examples

$$\text{Recall-} \quad \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{7}$$

e.g. Taking the example of confusion matrix of EHOG-PCA-NN (Fig. 9) to find the Recall refer formula no 7:

$$\text{Recall-} \quad \frac{39}{39+0} = 100\% \tag{8}$$

Sensitivity/True Positive Rate: It is defined as the correctly positive Classified Samples/examples is divided by the True Positive Samples/examples.

The calculation is same as calculated in recall.

$$\text{Sensitivity-} \quad \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{9}$$

4) **Specificity**: It is defined as the Specificity which is calculated as the number of correct negative predictions divided by the total number of negatives.

$$\text{Specificity-} \quad \frac{\text{TN}}{\text{FP}+\text{TN}} \tag{10}$$

$$1 - \text{ specificity/False Positive Rate} \quad \frac{\text{TN}}{\text{FP}+\text{TN}} \tag{11}$$

e.g. Taking the example of confusion matrix of EHOG-PCA-NN (Fig. 9) to find the Specificity refer formula no 10:

$$\text{Specificity-} \quad \frac{40}{1+40} \quad = 97.5\% \tag{12}$$

While classification of dataset through different feature extractor with different parameters have been calculated with the help of confusion matrix.

- **Calculation of Error Rate**

It is defined as the "number of incorrect predictions" that is divided by the "total number of the dataset".

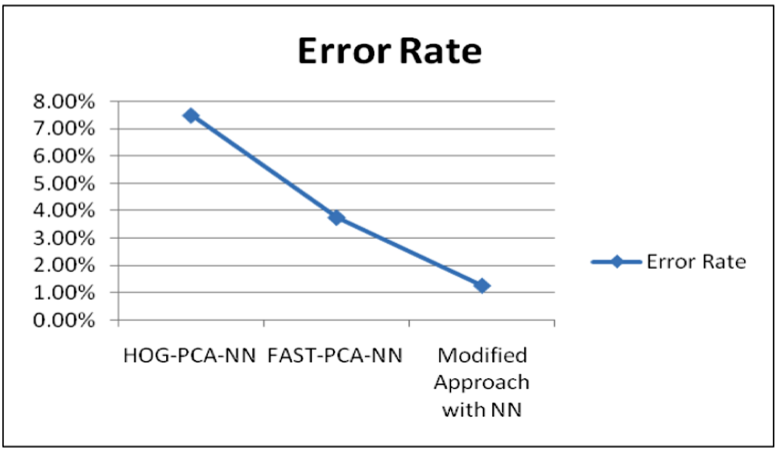


Fig. 10. Error rate between different techniques

In the above figure shows that the error rate in different Techniques with different Feature Extractor and that can be calculated by the Number of Incorrect Predictions divided by Total number of dataset.

$$\text{Error Rate:-} \quad \frac{FP + FN}{P + N} \tag{13}$$

Table 1. Error rate of different techniques

S. no	Technique/methods	Error rate
1	HOG-PCA-NN	7.5%
2	FAST-PCA-NN	3.75%
3	EHOG/modified approach with NN	1.25%

From the above Table 1 conclude that the EHOG/Modified Technique has less error rate as compared to other Techniques.

8.2 Analysis of Different Approaches with Different Parameters

In Table 1 shows the results analysis with different feature extractor with different parameters that have been calculated through confusion matrix. Confusion matrix will be vary with techniques. For classification neural network has been used with every feature extractor. However PCA (principal component analysis) has been used for dimensionality reduction.

Table 2. Analysis of Different Approaches with Different Parameters

Different approaches	Parameters				
	Accuracy	Specificity	Sensitivity	Precision	Recall
HOG-PCA-NN	92.5%	92.80%	92.10%	92.10%	92.10%
FAST-PCA-NN	96.3%	97.5%	97.3%	96.3%	94.3%
EHOG-PCA-NN	98.7%	100%	97.5%	100%	97.5%

Table 2 shows that the analysis of different parameters and the proposed feature extractor (EHOG) with the classifier which is Neural Network (NN) gives better result in every parameter. In terms of Accuracy (EHOG-PCA-NN) gives 98.7%, FAST-PCA-NN gives 96.3%, HOG-PCA-NN gives 92.5%. In precession (EHOG) gives 100%, FAST-PCA-NN gives 97.3%, HOG-PCA-NN gives 92.1%, In Recall (EHOGPCA-NN) gives 97.5%, FAST-PCA-NN gives 94.3%, HOGPCA-NN gives 92.5%, In Sensitivity(EHOG-PCA-NN) gives 97.5%, FAST-PCA-NN gives 97.3% and HOG-PCA-NN gives 92.10%. In specificity (EHOG-PCA-NN) gives 100%, FAST-PCA- NN gives 97.5%, HOG-PCA-NN gives 92.8%.

8.3 Analysis of F-Measure in Different Approaches

The F1 score measure the performance of the model and it is the harmonic average of the recall and the best value for the precision and recall is 1 and worst is 0.

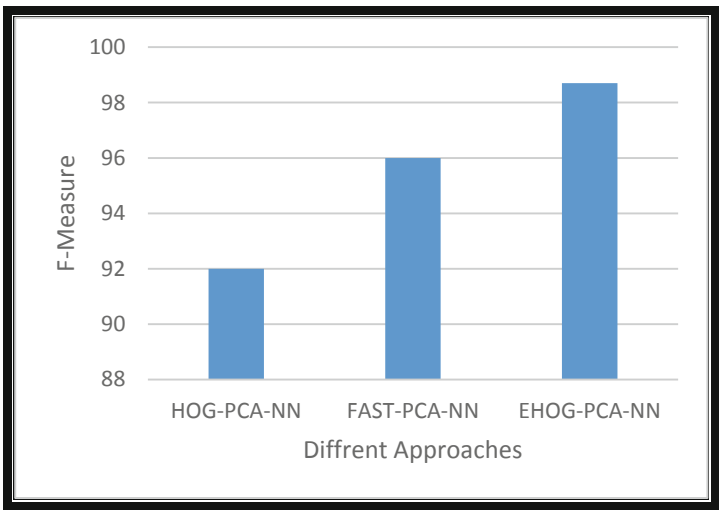


Fig. 11. Analysis of F-measure in different approaches

The Fig. 10 shows that the measure of a test's accuracy. It considers both the recall' r' and precision 'p' of the test to compute the test accuracy: p is the number of positive corrected results divided by the number of all positive results returned by the classifier, and r is the number of positive corrected results divided by the number of all relevant samples and sample should be considered as positive.

$$\mathrm{F1} = 2 * \frac{\mathrm{Precision} * \mathrm{recall}}{\mathrm{Precision} + \mathrm{recall}} \tag{14}$$

e.g. Taking the example of confusion matrix of EHOG-PCA-NN (Fig. 9) to find the F1 refer formula no. 14 (Table 2).

$$\begin{aligned} \mathrm{Precision} &= 97.5\% \qquad \mathrm{Recall} = 100\% \\ \mathrm{F_1} &= 2 * \frac{97.5 * 100}{97.5 + 100} = 98.73\% \end{aligned} \tag{15}$$

Hence, this section discusses simulation results with our proposed algorithms (discussed in Sect. 8). Also, it discuss the comparison/performance of our proposed algorithm with other existing work. Next section will conclude this work in brief with some future enhancements.

9 Conclusion

In this work proposed an Enhanced Histogram of Oriented Gradient (EHOG) for extraction of EUS (Epizootic ulcerative syndrome) features. The performance comparison of proposed EHOG with Neural Network gives better result in detection of EUS disease as compared to other feature extractor with different parameters as experimentation done in the paper. In near future, Internet of Thing (IoT) will be used in automated fish farming, agriculture in many other sectors for increasing productivity. Moreover Blockchain can also be used for real-time surveillance in the farming of the fish. EHOG with Neural Network can also be applied on large number of dataset in future. However other techniques or algorithm of deep learning can be used.

Acknowledgment. The EUS disease images of fish 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. kumar Pagrut, N., Ganguly, S., Jaiswal, V., Singh, C.: An overview on epizootic ulcerative syndrome of fishes in India: a comprehensive report. J. Entomol. Zool. Stud. **11**(4), 1941–1943 (2017)
2. Suresh, A.J., Asha, P.: Human action recognition in video using histogram of oriented gradient (HOG) features and probabilistic neural network (PNN). Int. J. Innov. Res. Comput. Commun. Eng. **4**(7), 13255–13263 (2016)
3. Burge, C.A., et al.: Climate change influences on marine infectious diseases: implications for management and society. Ann. Rev. Mar. Sci. **6**, 249–277 (2014)
4. Lafferty, K.D., et al.: Infectious diseases affect marine fisheries and aquaculture economics. Ann. Rev. Mar. Sci. **7**, 471–496 (2015)
5. Malik, S., Kumar, T., Sahoo, A.K.: Image processing techniques for identification of fish disease. In: IEEE 2nd International Conference on Signal and Image Processing (ICSIP), pp. 55–59 (2017)
6. Antony Seba1, P., Rama Subbu Laskhmi, S., Umamaheswari, P.: Fish recognition based on HOG feature extraction using SVM prediction. Int. J. Adv. Res. Comput. Commun. Eng**. 6** (5), 296–299 (2017)
7. Ansari, F.J.: Hand gesture recognition using fusion of SIFT and HoG with SVM as a classifier. Int. J. Eng. Technol. Sci. Res. IJETSR **4**(9), 206–210 (2017)
8. Maity, U., Mukherjee, J.: Automated color logo recognition technique using color and hog features. Int. J. Comput. Appl. **170**(2), 38–41 (2017)
9. Khan, H.A.: MCS HOG features and SVM based handwritten digit recognition system. J. Intell. Learn. Syst. Appl. **9**(2), 21–33 (2017)
10. Babri, U.M., Tanvir, M., Khurshid, K.: Feature based correspondence: a comparative study on image matching algorithms. Int. J. Adv. Comput. Sci. Appl. **7**(3), 206–210 (2016)
11. Dalal, N., Triggs, B., Schmid, C.: Human detection using oriented histograms of flow and appearance. In: ECCV 2006, pp. 428–441 (2010)

12. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), pp. 886–893 (2005)
13. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM **24**(6), 381–395 (1981)
14. Haiam, A., Abdul-Azim, H.A.: Human action recognition using trajectory based representation. Egypt. Inform. J. **16**(2), 187–198 (2015)
15. Pooja, G., Revansiddappa, S.K.: Abnormal activity detection using HOG features and SVM classifier. Int. J. Adv. Res. Electr. Electron. Instrum. Eng. **5**(4), 381–395 (2016)
16. Adetiba, E., Olugbara, O.O.: Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features. Sci. J. **2015**, 1–17 (2015)
17. Tian, S., Lu, S., Su, B., Tan, C.L.: Scene text recognition using co-occurrence of histogram of oriented gradients. In: Proceedings of the 2013 12th International Conference on Document Analysis and Recognition, Washington, DC, USA, pp. 912–916 (2013)
18. Bagum, N., Monir, M.S.: Present status of fish disease and economic losses due to incidence of disease in rural freshwater aquaculture. J. Innov. Dev. Strateg. (JIDS) **7**(3), 48–53 (2013)
19. Lyubchenko, V., Matarneh, R., Kobylin, O.: Digital image processing techniques for detection and diagnosis of fish diseases. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **6**(7), 79–83 (2016)
20. Kumar, V., Roy, S., Meena, D.K., Sarkar, U.K.: Application of probiotics in shrimp aquaculture: importance, mechanisms of action, and methods of administration. Rev. Fish. Sci. Aquac. **24**(4), 342–368 (2016)
21. Malik, S., Kumar, T., Sahoo, A.K.: A novel approach to fish disease diagnostic system based on machine learning. Adv. Image Video Process. **5**(1), 49–57 (2017)
22. ShavetaMalik, T.K.: Various edge detection techniques on different categories of fish. Int. J. Comput. Appl. **135**(7), 6–11 (2016)
23. Malik, S., Kumar, T., Sahoo, A.K.: Fish disease detection using HOG and FAST feature descriptor. Int. J. Comput. Sci. Inf. Secur. (IJCSIS) **15**(5), 216–221 (2017)
24. Cai, Z., Yu, P., Liang, Y., Lin, B., Huang, H.: SVM-KNN algorithm for image classification based on enhanced HOG feature. In: Proceedings of the International Conference on Intelligent systems and Image Processing (2016)
25. Stella, X.A., Sujatha, N.: Performance analysis of GFE, HOG and LBP feature extraction techniques using kNN classifier for oral cancer detection. J. Netw. Commun. Emerg. **6**(7), 50–56 (2016)
26. Niblack, W.: An Introduction to Digital Image Processing. Strandberg Publishing Company (1985)
27. Antony Seba, P., et al.: Fish recognition based on HOG feature extraction using SVM prediction. Int. J. Adv. Res. Comput. Commun. Eng. **6**(5), 296–299 (2017)
28. Santra, A.K., Christy, C.J.: Genetic algorithm and confusion matrix for document clustering. IJCSI Int. J. Comput. Sci. Issues **9**(1), 322 (2012)
29. Powers, D.M.W.: Evaluation: from precision, recall and F-measure to ROC: informedness, markendness & correlation. J. Mach. Learn. Technol. **2**(1), 37–63 (2011)
30. Lopes, J.N.S., Gonçalves, A.N.A., Fujimoto, R.Y., Carvalho, J.C.C.: Diagnosis of fish diseases using artificial neural networks. IJCSI Int. J. Comput. Sci. Issues **8**(6), 68–73 (2011)