

# Content-based Fish Classification Using Combination of Machine Learning Methods

**S.M. Mohidul Islam**

Computer Science and Engineering Discipline, Khulna University, Khulna, Bangladesh  
E-mail: mohid@cse.ku.ac.bd

**Suriya Islam Bani**

Computer Science and Engineering Discipline, Khulna University, Khulna, Bangladesh  
E-mail: suraiyaislambani@gmail.com

**Rupa Ghosh**

Computer Science and Engineering Discipline, Khulna University, Khulna, Bangladesh  
E-mail: rupa150218@gmail.com

Received: 24 April 2020; Revised: 11 June 2020; Accepted: 04 July 2020; Published: 08 February 2021

**Abstract:** Fish species recognition is an increasing demand to the field of fish ecology, fishing industry sector, fisheries survey applications, and other related concerns. Traditionally, concept-based fish species identification procedure is used. But it has some limitations. Content-based classification overcomes these problems. In this paper, a content-based fish recognition system based on the fusion of local features and global feature is proposed. For local features extraction from fish image, Local Binary Pattern (LBP), Speeded-Up Robust Feature (SURF), and Scale Invariant Feature Transform (SIFT) are used. To extract global feature from fish image, Color Coherence Vector (CCV) is used. Five popular machine learning models such as: Decision Tree, k-Nearest Neighbor (*k-NN*), Support Vector Machines (SVM), Naïve Bayes, and Artificial Neural Network (ANN) are used for fish species prediction. Finally, prediction decisions of the above machine learning models are combined to select the final fish class based on majority vote. The experiment is performed on a subset of 'QUT\_fish\_data' dataset containing 256 fish images of 21 classes and the result (accuracy 98.46%) shows that though the proposed method does not outperform all existing fish classification methods but it outperforms many existing methods and so, the method is a competitive alternative in this field.

**Index Terms:** Content, Global feature, Local texture, Combined Model, Fish species.

## 1. Introduction

According to some studies there are almost 1.2 million different types of fishes exist in the world but only 10 to 20 percent of these have discovered. Besides this, there exists many fishes that are not being classified yet and many fishes have extenuated without being classified. So, fish prediction is a mostly demand task in practical.

There are two methods to classify fish from image. They are concept or text-based methods and content-based methods. Concept-based fish identification analyzes the metadata associated with the fish image data. Having manually annotated fish in a large database can be time consuming, limit the scope of queries to the set of predetermined criteria, and it may not capture the keywords desired to describe the classes [1]. Sometimes same type or class fishes could have stored using different keywords or definitions which can create difficulties for user. That's why Text based process is not so efficient in case of classification [2]. On the other hand, content-based method is based on the appliance of computer vision techniques. Its intention is to retrieve the visual or low level information or features that correspond to the original image [1,3].

There are many existing methods to perform this task but many of them such as [4-5,17-20] shows less accuracy. In this research the main objectives are to classify fish from images using content-based image feature and combining the prediction decision of machine learning methods to get better recognition accuracy. Here both global and local features are used. CCV is used as color feature descriptor (global color) and LBP, SIFT, and SURF are used as texture feature descriptors (local texture) to create a fusion of both global and local features. For fish class prediction, five well known machine learning techniques are used and they are ANN, SVM, k-NN, Decision Tree and Naïve Bayes. Finally, the class/species of the query fish image is selected based on prediction of the majority. The method combines the five mentioned machine learning models to produce improved results than each of the used machine learning methods.

The rest of this paper is structured as follows: Section 2 describes some related works on Fish image processing. The proposed method of fish classification is explained in section 3. Section 4 presents the experiments and evaluation results. Conclusions and future works are given in the final section.

## 2. Literature Review

K. Anantharajah et al. [4] propose a local region based inter-session variability modeling approach. They apply it challenge real world fish image data. This Local inter-session variability approach provides a comparative performance improvement of, on average, 23%. It also affords a performance improvement of 35%.

In paper [5], deep neural networks (DNN) having different filter sizes and different numbers of convolutional layers are used for classifying fish. In their learning dataset, all of their tested structures were succeeded whereas in case of test dataset, the larger filters with less deep structures provide better results.

Paper in [17] proposed a fish classification methodology based on image segmentation, low level feature selection, and geometrical parameter techniques using decision tree and neural network. Their main contribution is enrichment recognize and classify fishes based on image and to develop a fish classification prototype. This method categorize the fish into its cluster and then to classify into non-poison or poison fish, and finally to classify the non-poison fish into its family.

In [18], the fish is taken in a meticulous environment with strong contrast and persistent illumination between fish object and its background. Using chain code method, it extracts shape features from segmented image. Their system can only detect the fish if both the trained and query images are of same size.

The aim of the work in [19] is to propose a fish identification system that functions in the underwater environment to understand fish behavior. Classification task is accomplished by combining texture and shape features. They used a database of ten different species and achieve an average accuracy of about 92%.

The paper presents in [20] proposed a Balance-Guaranteed Optimized Tree with Reject option (BGOTR)-based hierarchical fish classification method which identifies the 15 common fish species and detects new fish species in a non-constrained natural underwater environment. They apply a Gaussian Mixture Model and Bayes rule as a reject option after hierarchical classification. To do this, they estimate the posterior probability of being a certain fish species and then screen out less assertive decisions. They used a dataset from the south Taiwan sea and achieve noteworthy progresses compared to many contemporary techniques.

## 3. Materials and Methods

The proposed system involves two phases: Training phase and testing phase. The fig. 1 of the following demonstrates the architecture of the proposed system. In fig. 1, solid arrow line illustrates the data flow of the training phase and the dashed arrow line illustrates that of the testing phase. In training phase, after preprocessing and extracting feature from training images, each of the five machine learning models are trained using same feature vector data. In testing phase, the query image is preprocessed (resized) and its features are extracted in similar way of training phase. Then each trained machine learning models finds the fish class of the query image and finally select the final fish class (based on majority vote). In section A and B of the following, these phases are described in details.

### 3.1. Training Phase

In training phase, the machine learning models are trained using some known fish data. To accomplish this, the steps in the below are followed.

1) *Image Acquisition*: To acquire image, we split the dataset into 10 subsets and perform the training and testing 10 times. Each time 9 subsets of dataset are used for training and left one subset for testing.

2) *Image Resizing*: It is possible that all images size is not same, for that reason all images are resized to 400×400. Then the resized images are sent to the feature extraction process.

3) *Feature Extraction*: Fish features are machine understandable code that describes the characteristics of different fishes that make it easy to recognize the same types of fish image or images from a huge dataset.

There are some characteristics of fish images such as a wide variety of fish species, sheer numbers, head size, abdomen type, tail length, etc. The image features, both local and global features play vital role in fish recognition system. Different types of fishes have different color, texture, and shape features in their head, tail, and abdomen parts. Most of the existing works in fish recognition domain has been done using global or low level features. We see that fish classification system using local or high level features can generate improved results as compared to low level features. Both Color and texture feature are the most used spectacular features and mightiest features to describe fish. This paper considers the combination of global color feature and local texture features.

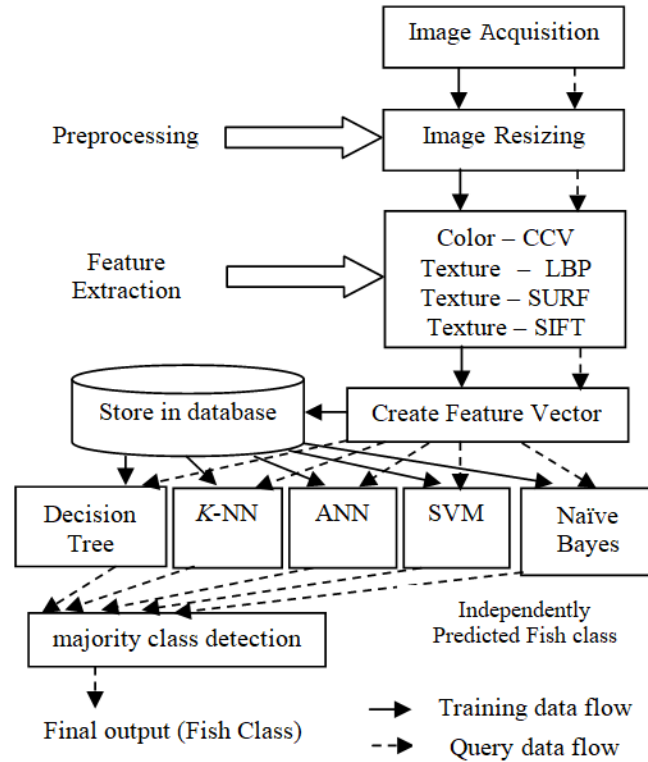


Fig.1. Proposed system architecture

To extract color feature of fish, we have used Color Coherence Vector (CCV). CCV classifies each pixel either coherent or incoherent where coherent pixels are part of a big connected component but incoherent pixels are not [6]. Feature extraction process of CCV is shown in the following:

Step 1: Blur the image.

Step 2: Discretize the color-space (image’s colors) into  $n$  distinct color.

Step 3: Need to categorize each pixel either as coherent or incoherent. It is computed as follows

- Find connected components for each discretized color.
- Calculate the value of ‘ $\tau$ ’ (specified value). Normally it is about 1% of image’s value.
- Any connected component with number of pixels more than or equal to ‘ $\tau$ ’ then its pixels are considered coherent and others are incoherent.

Step 4: Consider two values ( $C$  and  $N$ ) for each color, where,  $C$  = number of coherent pixels, and  $N$  = number of incoherent pixels. So number of pixels = the summation of all colors of  $C$  and  $N$ . Find the distance using following distance Function,

$$D(a, b) = \sum_{i=0}^n (|a_{C_i} - b_{C_i}|) + (|a_{N_i} - b_{N_i}|) \tag{1}$$

Where,  $C_i$ = number of coherent pixels colored with  $i$ ,  $N_i$ = number of incoherent pixels colored with  $i$ ;  $a$  and  $b$  represent  $C$  and  $N$ .

The steps of the above consider information about color spatial dispersal between pixels in its coherence component. The purpose of having a distance function is to enumerate the unlikeness between training and test images.

To extract local features of fish, we have used three different feature descriptors which have higher speed and accuracy: Speeded-Up Robust Features (SURF), Local Binary Pattern (LBP), and Scale Invariant Feature Transform (SIFT). We extract local texture features from head, abdomen and tail of fish using these descriptors.

SURF works on two stages: (1) key-point detection and (2) key-point description [10]. SURF uses primary images and effectual scale space construction for generating key-points and descriptors effectively [7]. Integral image is very important while detecting the interest points. Integral image can be found as follows:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \tag{2}$$

Here  $S(x,y)$  represents the addition of all pixels of input image  $I$  within a rectangular region initiated by the location  $(x,y)$  and the origin.

In the first stage, integral image helps the first computation of approximation of Laplacian of Gaussian (LoG) using a box filter. To identify the key points, Hessian Matrix criterion is calculated and used as a measure of local changes around a point. The point is chosen founded on the value of determinant. The point for which the determinant is maximum, is chosen. The Hessian matrix  $H(p, s)$  at point  $p(x, y)$  and scale  $s$  in an image is calculates as follows:

$$H(p, s) = \begin{pmatrix} L_{yx}(p, s)L_{xy}(p, s) \\ L_{yx}(p, s)L_{xy}(p, s) \end{pmatrix} \quad (3)$$

In second stage, each identified key points are allocated first to a specific orientation. The key point is centered on a square window and direction along based on previously acquired, the SURF descriptor is calculated. The window is partitioned into  $4 \times 4$  sub-regions. Then Haar wavelet of size  $2s$  is calculated within each sub-region, where  $s$  is the size of scale. Each sub-region provides 4 values. Thus, SURF provides 64D descriptors vectors. Finally, the values are normalized to unit length.

Uniform version of LBP is used here for fish classification. LBP features are extracted going through the following steps [8]:

Step 1: Have to divide the examined window or image into cells ( $16 \times 16$  pixels per cell).

Step 2: Compare each pixel in a cell with the surrounded 8 neighbor pixels.

Step 3: If the center pixel's value is greater than the neighbor's value, write 0, otherwise write 1. This gives an 8 digit binary number.

Step 4: Draw a histogram over the cell from which we can see for each combination which pixels are smaller and which are greater than the center. This histogram can be seen as a 256-dimensional feature vector.

Step 5: Do normalize the histogram.

Step 6: Concatenate the normalized histograms of all cells. By doing this, for entire window or image we get a feature vector.

SIFT is based on the Histogram of Gradients. Unlike SURF, SIFT produce too many key points among them some are unstable. Low contrast key points are discarded and points on edges are filtered among all the located points. Feature extraction process of SIFT [9,11] is shown in below:

Step 1: Create a  $16 \times 16$  window around the key point that is split into sixteen  $4 \times 4$  windows.

Step 2: Gradient magnitude and orientations are calculated for each  $4 \times 4$  window and orientations (in range 0-44 degree 45-89 degrees and so on) are put into 8 bin histogram. This is done on all sixteen  $4 \times 4$  windows.

Step 3: After normalizing all 16 windows, get a value  $16 \times 8 = 128$  numbers of feature vector for the key point.

4) *Create Feature Vector and Store into Database:* After extracting global colour feature and local texture features, a feature vector for each image is created and stored in the feature database.

### 3.2. Testing Phase

The image acquisition phase is similar to training phase except in every iteration of 10-fold cross validation only 1-fold of the dataset is kept leave for testing. Image resizing and feature extraction steps are similar to those of training phase. After extraction of features from each query image a feature vector is created and the feature vector is sent to the machine learning models for class prediction of fish.

1) *Machine Learning Models:* We have used five different machine learning models: Support Vector Machines (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbors (K-NN), Naïve Bayes, and Decision Tree.

K-NN is a non-parametric based classification algorithm where  $k$  represents the number of nearest neighbors [13]. This algorithm calculates  $k$  nearest neighbor and provides the decision based on majority. Euclidean distance is calculated for calculating result and the value of  $k$  taken is 5 in this research.

Naïve Bayes calculates the probability that a particular fish belongs in a certain class. For calculating probability for each class, Naïve Bayes classifier use Bayes' theorem [14]. Kernel distribution has been used to get better result.

SVM is a binary classifier [15]. SVM is used for classifying both linear and non-linear data. Kernel function is used to classify the non-linear data. Among all kernel functions Gaussian kernel function is used in this research.

Decision Tree is a top-down recursive tree induction algorithm which returns a fitted classification tree based on the input values. A leaf node of a tree indicates the value of the targeted class [14]. CART algorithm is used where Gini index measures the impurity of a set of training tuples.

As there are multiple classes Back Propagation Neural Network (BPNN) is used for this research work. In BPNN, each layer of the model learns some features. This algorithm is created for computing error deviation with a view to

hidden activities and updating weight accordingly [16]. In this research, a BPNN model is created where there are 12 hidden layers and 6 iterations has taken for creation of a network.

2) *Majority Class Detection*: It is a prediction model that is composed of a set of models, where each model predicts the output independently and final output is made by aggregating the outputs of these models. Majority vote based aggregating method is used here, where five different machine learning models independently predict fish class for the query image and finally the fish class is chosen based on majority [12]. If we compare this model with a single classifier, it is quite sure that combining model output based on majority has significantly greater accuracy.

#### 4. Experiment and Analysis

For experiments, QUT\_fish\_data dataset [21] is used. The dataset contains 4,405 fish images of 465 species of fish, captured in different environment and background conditions. Environment includes both sunny, cloudy, rainy day. Background includes both simple and complex background (background with another object). Some sample images of the dataset are shown in fig. 2.

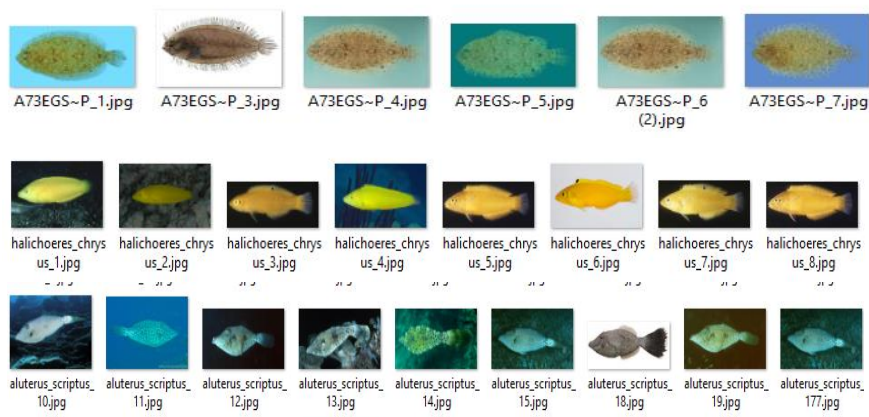


Fig.2. Sample images of QUT\_fish\_data dataset

For our experiment, we have used 256 fish images of 21 classes from the dataset. Fig. 3 shows GUI of our implementation using MATLAB tool. It shows the sample query image and image after preprocessing and then extracted the features (resized image). Feature vector is created and the vector is sent to the individual machine learning model for class prediction. Each model gives their decision. Output class is selected using prediction by majority models.

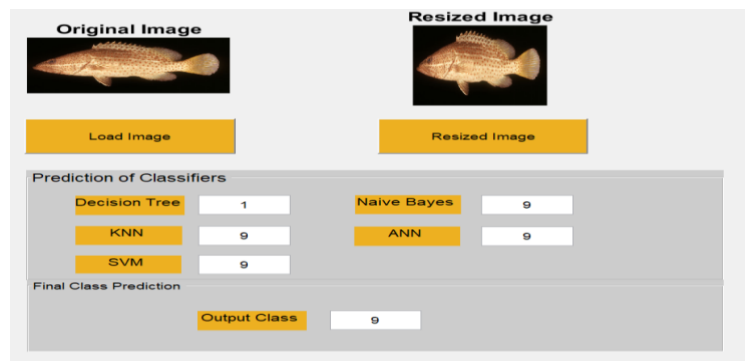


Fig.3. Class prediction for a fish

The experimental process used here is 10-fold cross validation. According to this process, for each round of 10-fold cross validation, 90% of fish images are placed for training and 10% are placed for testing. This iterates 10 times. After completing 10 rounds, final accuracy is calculated using aggregation. Table1 shows accuracy of each round and final accuracy for each feature descriptor.



Table 1. Accuracy using various features

	CCV	LBP	SIFT	SURF	CCV+ LBP+ SIFT+ SURF
Round 1	84.61%	84.61%	96.15%	100%	100%
Round 2	57.69%	80.77%	96.15%	96.15%	100%
Round 3	61.54%	92.31%	96.15%	96.15%	100%
Round 4	76.92%	88.46%	96.15%	92.31%	96.15%
Round 5	61.54%	88.46%	92.31%	100%	96.15%
Round 6	73.07%	80.77%	96.15%	100%	100%
Round 7	69.23%	76.92%	96.15%	92.31%	100%
Round 8	65.38%	88.46%	96.15%	96.15%	96.15%
Round 9	73.07%	80.77%	92.31%	96.15%	96.15%
Round 10	57.69%	76.92%	96.15%	100%	100%
Average Accuracy	68.07%	83.85%	95.38%	96.92%	98.46%

Here it is clear that when features individual performance is comparing then we can see that the performance of CCV is lowest among the four features, whereas the performance of SURF is highest which is more appreciating.

We see the combination of four features (CCV, LBP, SIFT and SURF) accuracy is 98.46%, higher than accuracy of each individual feature descriptor. Moreover, combining prediction decision of each of the supervised machine learning models that means counting majority vote of prediction of the machine learning models helps to improve the results than each single model.

## 5. Conclusion

Tremendous assortment of colors, variety of shapes and diversity of textures are the most outstanding features of fishes. The proposed method uses color and local texture features to classify different fishes using combination of machine learning approach. To classify fishes, five classifiers Decision Tree, KNN, SVM, Naïve Bayes and ANN are used. Multiplicity votes of classifiers are reckoned as final classification result of desired fish and the experimental results illustrate that the system delivers a promising result which shows the improvement in accuracy for using majority decision than using single learning model. To many extents, this would help the fish researcher to recognize fish species from image. This work can be extended by using more dataset and fish classes.

## References

- [1] F. Long, H. J. Zhang and D. D. Feng, Fundamentals of Content-based Image Retrieval, In Multimedia Information Retrieval and Management, D. Feng Eds, Springer, 2003.
- [2] A. Goyal and E. Walia, "Variants of Dense Descriptors and Zernike Moments as Features for Accurate Shape-based Image Retrieval," Signal, Image Video Processing, pp. 1–17, 2014.
- [3] A. Goyal and E. Walia, "An Analysis of Shape Based Image Retrieval Using Variants of Zernike Moments as Features," Int. J. Imaging Robot. [Formerly known as "International J. Imaging" (ISSN 0974-0627)] Vol. 7; Issue No. 1; Year 2012; Int. J. Imag. Robot. ISSN 2231–525X; Copyr. © 2012 by IJIR (CESER Publ., vol. 7, 2012.
- [4] K. Anantharajah, Z. Y. Ge, C. McCool, S. Denman, C. B. Fookes, P. Corke, D. W. Tjondronegoro, and S. Sridharan, "Local inter-session variability modelling for object classification," 2014.
- [5] M. Avic and M. Sarigul, "Comparison of Different Deep Structures for Fish Classification". In *International Journal of Computer Theory and Engineering*, Vol. 9, No. 5, October 2017.
- [6] <https://tarekmamdouh.wordpress.com/2013/08/19/image-retrieval-color-coherence-vector>.
- [7] Bay, Herbert, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. "Speeded-up robust features (SURF)." *Computer vision and image understanding* 110, no. 3 (2008): 346-359.
- [8] [https://en.wikipedia.org/wiki/Local\\_binary\\_patterns](https://en.wikipedia.org/wiki/Local_binary_patterns).
- [9] Lowe, David G. "Object recognition from local scale-invariant features." In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150-1157. Ieee, 1999.
- [10] Khan, Nabeel Younus, Brendan McCane, and Geoff Wyvill. "SIFT and SURF performance evaluation against various image deformations on benchmark dataset." In *2011 International Conference on Digital Image Computing: Techniques and Applications*, pp. 501-506. IEEE, 2011.
- [11] <http://aishack.in/tutorials/sift-scale-invariant-feature-transform-features/>.
- [12] <https://people.cs.pitt.edu/~milos/courses/cs2750Spring04/lectures/class23.pdf>.
- [13] [https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm).
- [14] Jaiwei Han, Micheline Kamber, Jian Pei, "Data Mining Concepts and Techniques", Morgan Kaufmann, Third Edition.
- [15] <http://www.statsoft.com/textbook/support-vector-machines>.
- [16] Christopher M. "Bishop Pattern Recognition and Machine Learning", 2006.
- [17] Mutasem Khalil Sari Alsmadi, Prof. Dr. Khairuddin Bin Omar, Prof. Dr. Shahrul Azman Noah and Ibrahim Almarashdah, "Fish

recognition based on the combination between robust features selection, image segmentation and geometrical parameters techniques using artificial neural network and decision tree”, (IJCSIS) International Journal of Computer Science and Information Security, Vol.6, No. 2, 2009.

- [18] Kho Geok Hond, “Fish species recognition system based on chain code representation of shapes”, Universiti Teknologi Malaysia, May 19, 2011.
- [19] C. Spampinato, D. Giordano, R. Di Salvo, Y.-H. J. Chen-Burger, R. B. Fisher, and G. Nadarajan, "Automatic fish classification for underwater species behavior understanding," in *1st ACM International Workshop on ARTEMIS*, 2010, pp. 45-50.
- [20] P. X. Huang, B. J. Boom, and R. B. Fisher, "Hierarchical classification with reject option for live fish recognition," *Machine Vision and Applications*, vol. 26, no. 1, pp. 89-102, 2014.
- [21] <http://wiki.qut.edu.au/display/cyphy/Fish+Dataset>.

### Authors' Profiles



**S.M. Mohidul Islam** is a faculty member of the Computer Science and Engineering Discipline, Khulna University, Bangladesh. He received both B.Sc. Engg. and M.Sc. Engg. Degree from this University. Mr. Islam's research interests include Community ICT (ICT used for welfare of the human community), Machine learning, and Image Data Mining.



**Suriya Islam Bani** is a former student of the Computer Science and Engineering Discipline, Khulna University, Bangladesh. Her research interests include Machine learning, Digital Image Processing, and Artificial Intelligence.



**Rupa Ghosh** is a former student of the Computer Science and Engineering Discipline, Khulna University, Bangladesh. Her research interests include Computer Vision, Pattern Recognition, and Machine learning.

**How to cite this paper:** S.M. Mohidul Islam, Suriya Islam Bani, Rupa Ghosh, "Content-based Fish Classification Using Combination of Machine Learning Methods", International Journal of Information Technology and Computer Science(IJITCS), Vol.13, No.1, pp.62-68, 2021. DOI: 10.5815/ijitcs.2021.01.05