



Research Article

Machine vision based local fish recognition

Israt Sharmin¹  · Nuzhat Farzana Islam¹ · Israt Jahan¹ · Tasnem Ahmed Joye¹ · Md. Riazur Rahman¹ · Md. Tarek Habib¹

Received: 10 June 2019 / Accepted: 24 October 2019 / Published online: 1 November 2019
© Springer Nature Switzerland AG 2019

Abstract

Bangladesh has its own abundance of water resources which helps to identify its customs that are related to freshwater fish. Due to environmental issues along with some other reasons, the amount of water resources of Bangladesh is reducing day-by-day. Consequently, many of our territorial freshwater fishes are getting abolished. Thus, the new generation people of Bangladesh lacks the knowledge of local freshwater fish. For this problem, a solution has been found with the collaboration of vision-based technology. As a solution, a machine-vision based local freshwater fish recognition system is presented that can be proceed with an image of fish captured with a mobile or handheld device and recognize the fish in order to introduce the fish. To demonstrate the utility of the proposed expert system, several experiments are performed. At first, a set of fourteen features, which consists of four types of features, are presented. Then the color image has been converted into gray-scale image and the gray-scale histogram is formed. Image segmentation takes place using histogram-based method and then the features are extracted. PCA is used for decreasing the feature numbers. Three classifiers are used for recognizing fish, where SVM gives the highest accuracy showing a value of 94.2%.

Keywords Bangladeshi local fish · Machine vision · Thresholding · Principle component analysis · Support vector machine · Classifiers · *k*-nearest neighbor

1 Introduction

Bangladesh is well known in the world as a riverine country. Due to 700 rivers and embranchments, the gross domestic product (GDP) rate (3.69% in the fiscal year 2015–16) of the fisheries performs significant role in national economic sector [1]. Moreover, Bangladesh earns more than 2% from the inland fisheries sector in international trade in the fiscal year 2015–16 [1]. Fish is a great source of protein and Bangladeshis are used to have some fish in their daily meal. According to the current state of our country, some local freshwater fish is becoming extinct day-by-day because of environmental and some other reasons. River erosion and water pollution have become big threats for freshwater fish. That is why, most people of

the new generation do not know these fish. If things go on this way, the next generation will have no idea about these fish. Artamim, Anju, Arwari, Baim, Botika, Boumach, Bele, Chapila, Kholshes, Pabda, Punt, Meni, Bilchuri, Murari, Darkina some remarkable fish among them.

In this paper, an experimental research on local freshwater fish recognition is executed using machine-vision based approach. We want to introduce a method that can proceed the fish image and recognize which fish is in the image. A set of feature of fourteen features has been presented for recognizing the fishes. After segmenting the image, these features are extracted in order to form the feature vector. To reduce the dimensions of the feature vector, principal component analysis (PCA) is used. Three classifiers, namely support vector machine (SVM),

✉ Israt Sharmin, ima.sharmin23@gmail.com; Nuzhat Farzana Islam, nuzhat15-5316@diu.edu.bd; Israt Jahan, israt15-5461@diu.edu.bd; Tasnem Ahmed Joye, ahmed15-4721@diu.edu.bd; Md. Riazur Rahman, rrahman.diu@gmail.com; Md. Tarek Habib, md.tarekhabib@yahoo.com | ¹Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh.



k-nearest neighbor (k-NN) and ensemble are deployed for the classification of the fish.

The rest of the paper is organized as follows. Section 2 describes the current state of a solution to address the problem of automated fish recognition. Section 3 describes the entire methodology of our research work. Section 4 explains the fish as well as their features. Feature extraction and feature selection methods are also enlisted in this section. Section 5 shows the results and discussions of the research. Section 6 describes comparative result. Finally, summary, limitations and future scopes are presented in Sect. 7.

2 Literature review

In the sector of fish recognition, few works are done among them fish from Bangladesh are hardly found. On the other hand, there are adequate works for fruit recognition by their many characteristics. Like on their color and texture [2], grading system [3], color characterization [4], classification [5], detection of defective apple [6] and many works can be done by using machine-vision. Apart from this some works of food done by using machine-vision applications in recognition and aquatic food [7, 8], deep convolutional network with pre-training and fine-tuning [9]. It is a matter of woe that very few works have performed at fish sector.

In paper [10], some features of fish are measured in the system. The widths and heights of fish are counted at various location using a camera perpendicular. They used MAXSCAN frame grabber, ROISTORE video memory and MAXGRAPH display unit. For generating and training the neural network, they used a platform called VME system which is connected to a workstation. After training this network, 95% of the fish of different species were classified correctly. In paper [11], they proposed two-dimensional function that is used to identify as the intensity of image. Two-dimensional function of electromagnetic spectrum is used. They used computer software for transmitting electronic signals and images are captured by color machine-vision systems. Illumination, segmentation, feature extraction, classification/matching are performed. The accuracy of grey Mullet, St. Peter fish and Carp was indicated to be 100%, 92% and 89%. They used ANN (Artificial Neural Network) and got 99.8% authenticity for seven categories. The whole trainable system is capable to determine fish weight by using image. Image processing used for rigor mortis by automatic imaging-based system. In [12], they had given a method for classifying species of fish emerged on color and texture features and used a multiclass support vector machine (MSVM). In their work, 1024 × 768-pixel images of fish were captured. An automatic cropping program

was used to crop some area of the actual fish images and acquire 512 × 512-pixel images. Fish images were converted into HSV space and total six color features were extracted. Gray scale Histogram (GH) and Gray Level Co-occurrence Matrices (GLCMs) were used for texture analysis. Two types of OAO (One-Against-One) based on MSVMs were constructed and compared to serve for the classification of fish species. Acyclic graph MSVM(DAGMSVM) and voting based MSVM(VBMSVM) were designed and utilized in their study. These constructed MSVMs are also compared with the LIBSVM which is a nonlinear SVM software based on OAO algorithm. In [13], they had claimed that fish attributes can be computed and collected by size and shape measurements tools like size of mouth, angle of head, caudal fin length, dorsal fin length, caudal angle and the angle between the mouth and the eye through the distance and geometrical tools and the process of doing this kind of calculating is called feature extraction. Here using landmark point detection where we found that the one point to another point measure of fish. The percentage of accuracy was 89%. Researchers used 500 fish images for testing their method where 350 fish images for training and the rest for testing. They also need more data set for their work.

From above description, we can dissolve that none of any research work has performed on our local fish in which fish carry kudos of our Bangladesh. Moreover, most of the works involve fruit, food or sea fish's edge detection and other categories of recognition which are not up to the mark. The research gap between current states of the works can be viewed. Then there is not enough resources regarding this topic where we can discuss about other problems which are related with it.

3 Research methodology

A color image of a local freshwater fish is needed for starting the work of the system. As fish images were captured from local fish market, at that time fish were found dead. So the work is limited to fixed number of orientation. In this project, Traditional machine learning has been used over deep learning. It can be assume that the ideal model complexity produces the lowest generalization error [14]. And the classifiers that have been used give ideal model complexity. For using deep learning, huge image data needed but huge amount of image data cannot be collected in this project. So, traditional machine learning has been preferred in this work.

First of all, the fish image is altered into a 512 × 512 dimensional image which is fixed. After resizing, the color image is needed to convert into a gray-scale image, where the level of gray-scale is 256. If r, g and b represent as well

as red, green and blue values of a pixel in a color image of RGB and y illustrates as a value of gray-scale in the same pixel, according to [14], y can be written as:

$$y = 0.33 \times r + 0.56 \times g + 0.11 \times b \tag{1}$$

Then the image is segmented using one of the histogram-based methods, namely histogram peak technique [15]. It is one of the easy and mostly used for segmentation method. Two values of thresholding are calculated in order to convert the gray-scale image into a binary. The two thresholding values are known as θ_L and θ_H and these values are adaptive. These adaptation depends on images and values can be found by computing each corresponding to distinct section of image. The image, whose each pixel is illustrated by $p(x, y)$, is converted into a binary image, whose each pixel is illustrated by $bi(x, y)$, where ,

$$bi(x, y) = \begin{cases} 1, & \text{if } \theta_L \leq p(x, y) \leq \theta_H. \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The binary image contains two things—object and background. After the image is segmented, then the features have been extracted. Color intensity has been used for extracting feature as it is advantageous. For getting RGB mean value, the original image which is saved in buffer has been used for calculating RGB mean value. After getting mean value of RGB image. Then RGB image is converted into HSV for calculating mean of HSV value. As we know HSV is better than RGB for color analysis and is advantageous for feature extraction. Here, Hue (H) means the attribute of visual sensation which corresponds to color perception associated with the dominant colors and its value runs from 0 to 360°. Saturation (S) determines the degree of strength or purity where purity indicates how much white is added to the color and saturation ranges from 0 to 1. Value (V) measures the brightness of a color [16]. It can be stated from [16] that

$$H_1 = \cos^{-1} \frac{\frac{1}{2}[(R - G)(R - B)]}{\sqrt{(R - B)(G - B)(R - G)^2}} \tag{3}$$

$$S = 1 - \frac{3}{(R + G + B)} [\min RGB] \tag{4}$$

$$V = \frac{1}{3}(R + G + B) \tag{5}$$

After getting value for color intensity, other feature extractions have been proceed using segmented image. In total, fourteen features are extracted in this work. Fourteen features are too large to classify six classes of fish. And feature selection is best idea than comparing extracted feature. The curse of dimensionality is a situation that is observed

to happen in high dimensional space during analyzing and also organizing data, but it is no seen in low dimensional space as like third dimensional physical space [14]. In this project, at first the work is done with 14 features. Due to risk of curse of dimensionality, this features is reducing to 7 instead of 14 features. With the help of PCA system the process completed, as the ultimate goal of PCA (Principle Component Analysis) is to discover new dimension set which can capture better the variability of data. As before state that the curse of dimensionality mostly happens if the design data-set is higher than the range from 2 to 1000 [17]. As this research work has been proposed for mobile application, it will be favorable if there is data simplicity, computational and cost reducing. On account of this, PCA has been used in this research project. On the other hand, it cannot be reducing less than 7 features, because of fear of accuracy problem. So, feature selection is performed and PCA is reduces fourteen features into seven features. In next section, the features and feature selection have been described in details. The architecture of Machine Vision based expert system for recognizing fish has been shown in Fig. 1.

These feature vectors are feed, who are thus collected, to SVM, k-NN and Ensemble. After the classifiers are used on the trained set of data, the performance of the classifier can be counted using the set of test data. Only accuracy may not be good enough for evaluating classification when the performance is analyzed. So, there are considered other metrics required for evaluating classification performance. The number of false positives (FPs), false negatives (FNs), true positives (TPs), and true negatives (TNs) are reported to be A confusion matrix for a binary, i.e. 2-class problem [18]. In multiclass confusion matrix, the matrix will be of dimension $n \times n$ ($n > 2$), which lead itself to contain n rows, n columns and $n \times n$ entries in total. According to multiclass matrix described in [18], the values of FPs, FN, TPs and TNs for class $i = 1, 2, 3, \dots, n$ are calculated as,

$$TP_i = a_{ii} \tag{6}$$

$$FP_i = \sum_{j=1, j \neq i}^n a_{ji} \tag{7}$$

$$FN_i = \sum_{j=1, j \neq i}^n a_{ij} \tag{8}$$

$$TN_i = \sum_{j=1, j \neq i}^n \sum_{k=1, k \neq i}^n a_{jk} \tag{9}$$

After the approach described above works, the final confusion matrix carries the average values of the n confusion

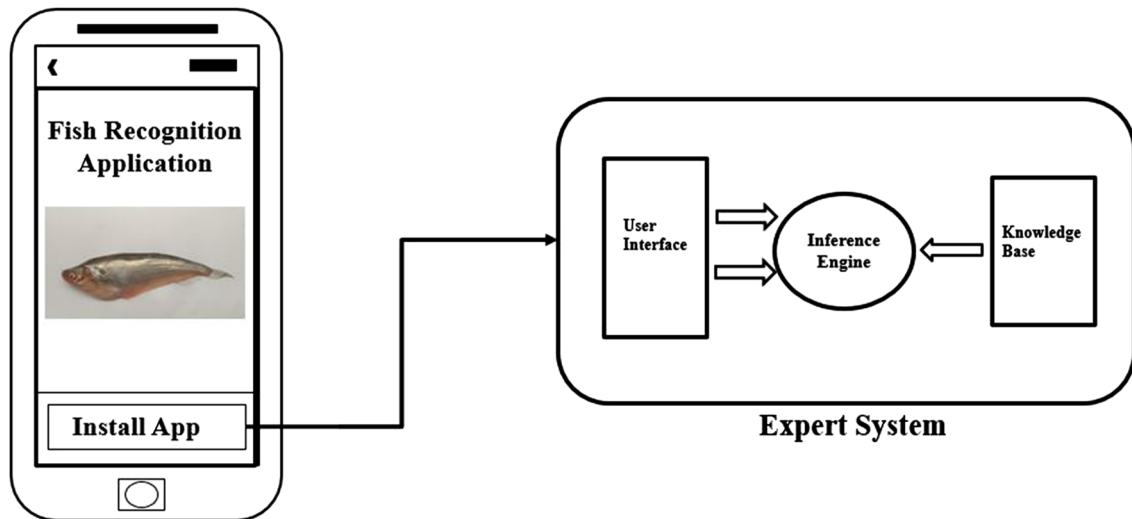


Fig. 1 The architecture of machine vision-based expert system for recognizing fish

matrices for each class and is of dimension 2×2 . Using this confusion matrix, accuracy, precision, sensitivity, false negative rate (FNR) and false positive rate (FPR) are calculated in percentage as [14]:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100 \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{11}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \tag{12}$$

$$FNR = \frac{FN}{TP + FN} \times 100 \tag{13}$$

$$FPR = \frac{Fp}{TP + FN} \times 100 \tag{14}$$

Here, TP, FP, FN, TN are the average of TP_i, FP_i, FN_i, TN_i . So, in this way, the three classifiers' performances are computed regarding these metrics and they become ready to classify fish images. Receiver operating characteristic (ROC) curves are also applied so that one can compare the performances of classifiers [14]. Here, three classifiers are applied for comparing the performance. The entire approach is shown with the block diagram in Fig. 2.

Fig. 2 Approach for recognition of local freshwater fish

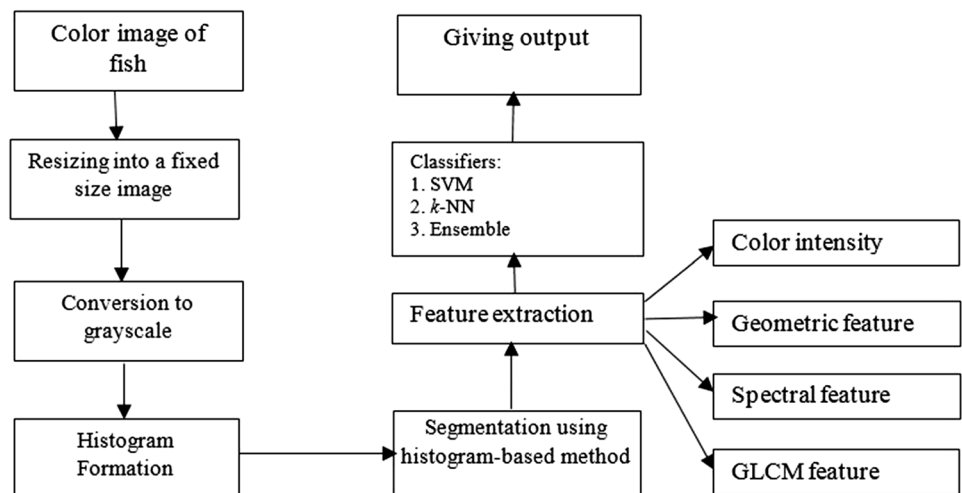
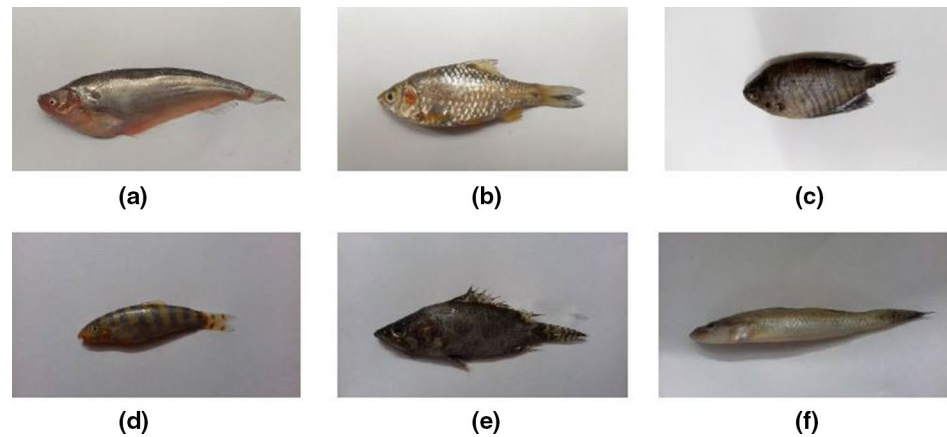


Fig. 3 Local fish that only find in Bangladesh: **a** Pabda (*Ompok pabda*), **b** Puti (*Puntius sophore*), **c** Kholse (*Awaous guamensis*), **d** Boumach (*Botia dario*), **e** Meni (*Nandus nandus*), **f** Bele (*Awaous guamensis*)



4 Local fish and feature description

4.1 Description of local fish

Fish observation is a very important part of our work, because it helps select features and recognize a local fish. In this work, we deal with six species of fish from all local freshwater fish, namely Pabda (*Ompok pabda*), Puti (*Puntius sophore*), Kholse (*Colisafas ciata*), Boumach (*Botia dario*), Meni (*Nandus nandus*), Bele (*Awaous guamensis*). They are shown in Fig. 3. Here is to mention that these fish are seldom found in local fish market which is not forbidden by the Government of Bangladesh. These fish differ in color, shape, size and texture. So, the features need to be selected by taking these into account.

4.2 Description of data

In this paper, an experimental research on local freshwater fish recognition is executed using machine-vision based approach. It has already mention in the previous section that the species of fish, we selected for this research are seldom found in local fish market which is not prohibited by the Government of Bangladesh. Images of fish are taken from different local fish markets. Capturing images of fish are the main source of collecting data. Pictures of fish have been taken using cellphone with high to low resolution and different angles. All six classes of fish images in different angle and in low resolution are shown in Fig. 4. Six species of fish helped us to our machine vision work. 180 pictures are taken of those fish species. Class imbalanced data-sets happen in many real-world applications where the class distributions of data are imbalanced [14]. As, some fish are not available when the research has been done, class imbalanced has been occurred. As for that same number of image cannot be used for each type of

Table 1 Number of fishes used

Fish name	Number of images
Puti	28
Pabda	30
Bele	31
Boumach	29
Meni	32
Kholshe	30

image in this work. Number of fish which are used in for work have shown in Table 1.

4.3 Feature extraction

Based on the observation of the six species of fish, a set of fourteen features is selected for the task of classification. The feature set consists of four types of features like color intensity, geometric, spectral and GLCM [12–14]. The features are extracted using the whole image and there is no need of cropping image or giving point. Here, the features are discussed in detail here.

4.3.1 Color intensity

For selecting features, the main priority is to consider color model at very first. The original image has been saved in buffer. Those saved image has been used for calculating RGB and HSV mean value, because in this work, background has been white and segmented image is not necessary. RGB and HSV color models are very advantageous. For numeric representation a color in the RGB color model is described by pointing how much of each of the red, green, and blue is included. Each color individually indicates zero to maximum values. This is the equation, we used to calculate red component, green component, blue component and mean of RGB space (μ_{RGB}) If the number

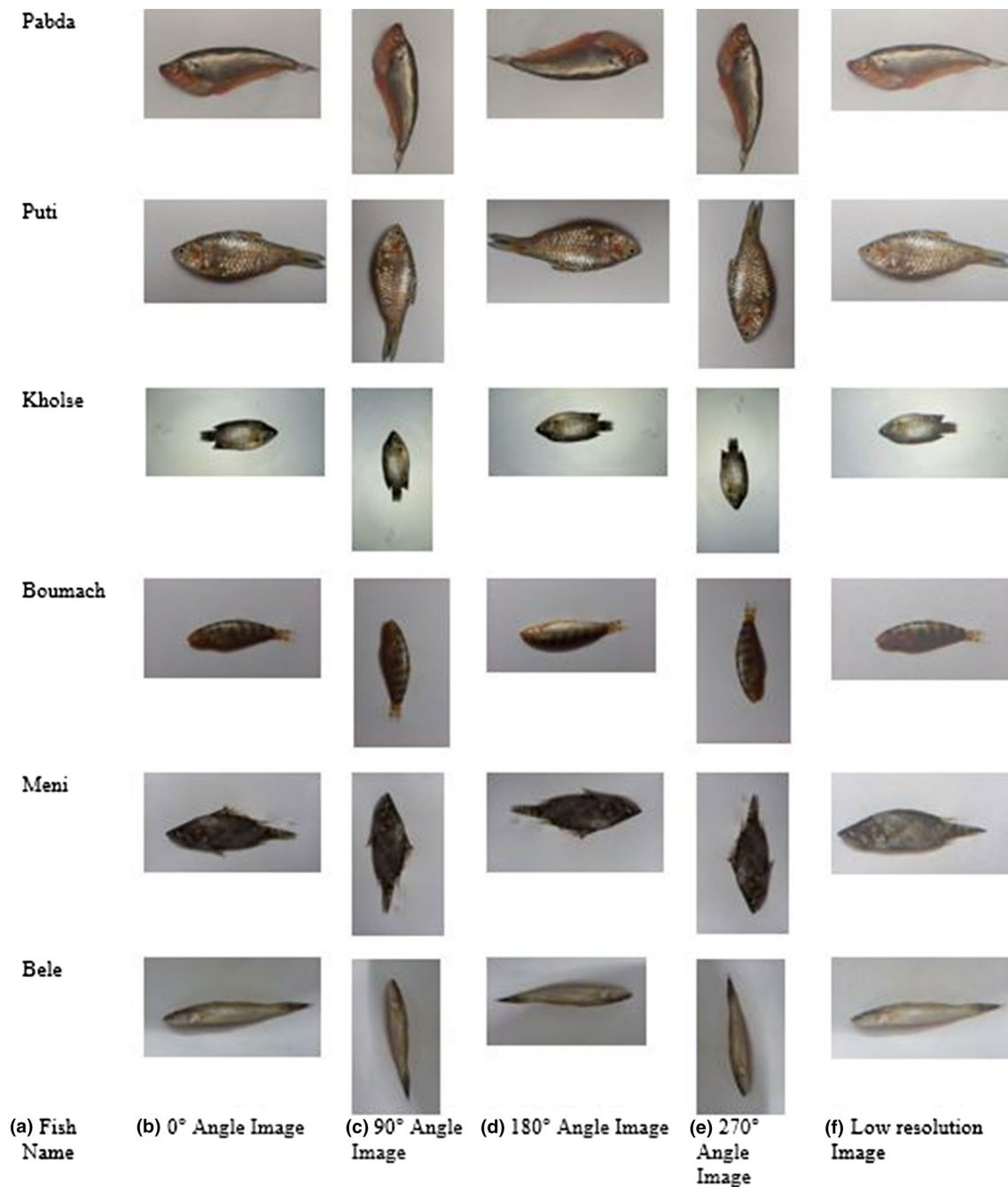


Fig. 4 **a** Fish name, **b** 0° angle image, **c** 90° angle image, **d** 180° angle image, **e** 270° angle image, **f** low resolution image

of calculated pixels is referred as N_{FI} in color images of fish (FI), as per [12], RGB mean becomes

$$\mu_{RGB} = \frac{1}{N_{FI}} \sum_{(i,j) \in FI} I(i,j) \tag{15}$$

where $I(i, j)$ is the value of the RGB pixel at location (i, j) . But RGB is not very efficient for color analysis. So, converting RGB image into HSV space is the alternative color

representation of RGB model where human vision perceives color attributes to make it more closely classify. Mean of HSV space (μ_{HSV}) refers to mean of hue, saturation, and brightness components of HSV space and N_{FI} refers to the number of calculated pixels in color images of fish (FI), as per [12], HSV mean becomes,

$$\mu_{HSV} = \frac{1}{N_{FI}} \sum_{(i,j) \in FI} I(i,j) \tag{16}$$

where $I(i, j)$ is the value of the HSV pixel at location (i, j) .

RGB mean and HSV mean, these two features are used as color intensity features.

4.3.2 Geometric

After segmentation, the following six geometric features are calculated.

1. Height and width: expressed in terms of number of pixels.
2. Area: expressed in terms of number of pixels.
3. Solidity: expressed the proportion of pixels in the set of convex.
4. Convex: the number of pixels in the convex hull.
5. Perimeter: the distance around the boundary of the object.
6. Mean Intensity: referred to a global measure of image.

4.3.3 Spectral

Fourier transformation is used for extracting feature. Fourier transform represents a sum of complex interpretive of different magnitudes, frequencies and phases of an image and DFT (Discrete Fourier Transform) is the sampled Fourier Transform which contain a set of sample that describe fully spatial image. The number of frequencies assembles the number of pixels in spatial domain image. Two dimensional DFT is given by [19],

$$F(k, l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-2\pi i (\frac{ki}{N} + \frac{lj}{N})} \tag{17}$$

Where, $f(i, j)$ is the image in spatial domain and the value of each point $F(k, l)$ is obtained by multiplying spatial image with assembling base function and summing the result. In this study, $F(0, 0)$ illustrates average brightness and $F(N - 1, N - 1)$ illustrates highest frequency [19].

4.3.4 GLCM

Several GLCM features are also used, which had been first suggested by Haralick et al. [20]. Suppose that $f(x, y)$ be a two-dimensional image of size $M \times N$ pixels with number of gray levels L . We can also presume that (x_1, y_1) and (x_2, y_2) are two pixels in $f(x, y)$, θ is the angle between the two and the ordinate and the distance of two pixel is d . Then a GLCM $P(i, j, d, \theta)$ becomes as per [20]:

$$P(i, j, d, \theta) = |(x_1, y_1), (x_2, y_2) \in M \times N : d, \theta, f(x_1, y_1) = i, f(x_2, y_2) = j| \tag{18}$$

In this study, contrast (C), correlation (ρ), energy (E), entropy (S) and homogeneity (H) are used as GLCM features. (19)–(23) are formulas for computing these five features:

$$Contrast : C = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - j)^2 P(i, j) \tag{19}$$

$$Correlation : \rho = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i, j, P(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{20}$$

$$Energy : E = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j)^2 \tag{21}$$

$$Entropy : S = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j) \log_2 P(i, j) \tag{22}$$

$$Homogeneity : H = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{P(i, j)}{i + (i - j)^2} \tag{23}$$

where μ_x, μ_y, σ_x and σ_y are the sum of expected and variance values for the row and column matrix entries, respectively. And $P(i, j)$ is the value of the pixel at location (i, j) .

4.4 Feature reduction

PCA for feature selection is used here. The objective of feature selection is selecting the features which grant for an exact description and reduce the dimensionality of input features [21]. In this work, there are fourteen features in total, which decrease into seven features using PCA as feature selection. Usually, the PCA technique translates n vectors $(x_1, x_2, \dots, x_i, \dots, x_n)$ from d -dimensional space to n vectors $(x'_1, x'_2, \dots, x'_n)$ in a new, d' -dimensional space as [21],

$$x'_i = \sum_{k=1}^{d'} a_{k,i} e_k, d' \leq d \tag{24}$$

where e_k is eigenvector and $a_{k,i}$ is the projections of original vectors x_i for $i = 1, 2, 3 \dots, n$ [21].

5 Experimental evaluation

Investigative experiment has been performed where machine-vision-based local fish recognition are followed as per Fig. 2. Capturing an image of local freshwater fish

is our first work to do. Then the captured image is altered into 512×512 dimensional image. Then the mean of RGB and HSV color will be calculated. Then the converted image is segmented by using histogram peak technique [22], where fish is separated from the background. All six classes of segmented fish images are shown in Fig. 5. After segmentation, fourteen features are extracted. By applying PCA on these fourteen features, seven features are come up with. Thus, a feature vector or input vector is formed. Here is to mention that classifiers are also performed for fourteen features but there is no significant

change in this work. One hundred and eighty color images of six different classes of local freshwater fish are taken for our work. So, the considered size of data set becomes 180. The entire set of data has been divided into two section known as testing and training. Hold-out method [14] has been applied randomly for electing the proportion of data which are reserved for testing and training. Sample set of data is divided into 60% as training set (96 images) and 40% as testing set (84 images). For avoiding model over fitting problem (that means to have low training error and low generalization

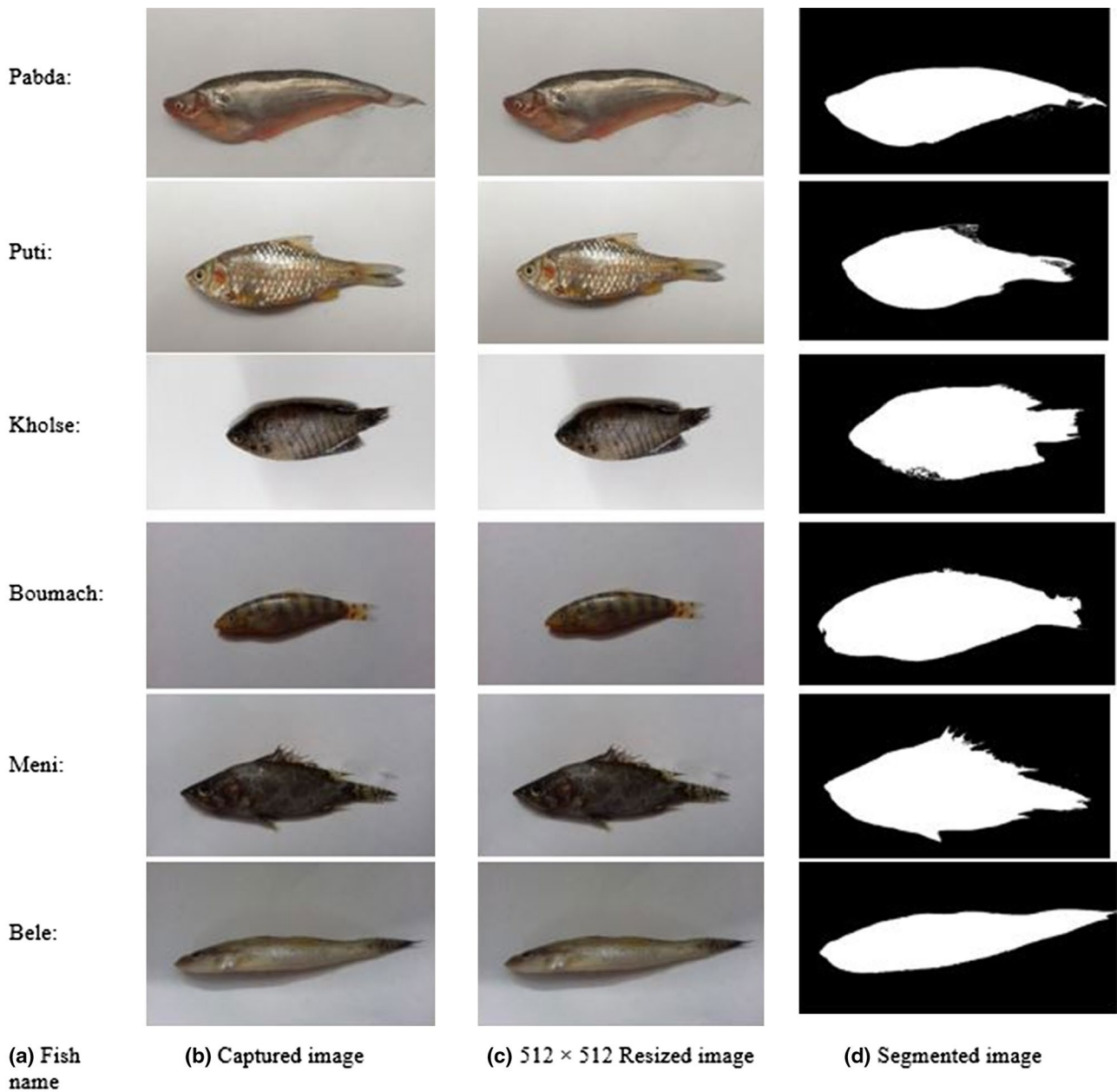


Fig. 5 **a** Fish name, **b** captured image **c** 512×512 resized image, **d** segmented image

Table 2 Detailed parameter specifications of the used three classifiers

Classifier	Specification
SVM	Kernel: linear C = 25,007
k-NN	Distance metric: euclidean distance k = 1
Ensemble	Type: bagged trees Size of each bag (percentage of size of training data) = 100% Trees maximum depth = unlimited Number of attributes that are randomly chosen = 0

error), a validation set has been applied [22]. According to this approach, the actual set is spliced of training data into two smaller subsets which are used for validation and training. 60% (58 images) of the training data set are considered for classifier building while the remaining 40% (38 images) is used for error estimation. As for finding the final trained classifier, holdout method is

repeated thrice. Using the test set, the performance is found out of the classifier after each implementation of holdout method. To build multiclass confusion matrix, the average is calculated of thrice-found results.

Then it comes to the matter of applying classifiers. Three classifiers are applied, namely SVM, k-NN and Ensemble for our work. SVM (support-vector machine) is a model of supervised learning that can analyze data used for classification and regression analysis and linear SVM has been used in this project [14]. As for, k-NN (k-nearest neighbors) is a non-parametric method which is used for pattern recognition. It is also used for classification and regression but in both cases, the input depends on the consistency of k closest training examples whenever the output depends on whether k-NN is used for classification or regression. And in this work k value is 1 [14]. Ensemble combines some base models in order to produce one optimal predictive model. For this work, Bagged Trees type has been used [14]. All the parameter of classifiers are come up with optimized value and those value are specific. Detailed parameter specifications of these three classifiers are listed in Table 2.

Table 3 Confusion matrix of SVM classifier

Predicted/true class	Pabda	Puti	Kholse	Boumach	Meni	Bele
Pabda	12	0	0	0	1	1
Puti	2	12	0	0	0	0
Kholse	1	1	12	0	0	0
Boumach	0	0	0	14	0	0
Meni	1	0	0	0	13	0
Bele	0	0	0	0	0	14

Table 4 Confusion matrix of k-NN classifier

Predicted/true class	Pabda	Puti	Kholse	Boumach	Meni	Bele
Pabda	12	1	0	0	1	0
Puti	0	13	0	1	0	0
Kholse	1	2	10	0	1	0
Boumach	0	0	0	14	0	0
Meni	0	1	1	0	12	0
Bele	2	0	0	0	0	12

Table 5 Confusion matrix of ensemble

Predicted/true class	Pabda	Puti	Kholse	Boumach	Meni	Bele
Pabda	12	1	0	0	0	1
Puti	1	13	0	0	0	0
Kholse	0	4	10	0	0	0
Boumach	0	0	0	14	0	0
Meni	2	2	1	0	9	0
Bele	0	0	0	0	0	14

Table 6 Comparison of the three experimentally evaluated classifiers

Classifier	Accuracy (%)	Sensitivity (%)	Precision (%)	False negative rate (%)
SVM	94.2	94.59	93	11.67
k-NN	92.6	93	91	15
Ensemble	88.4	87.33	89.4	23

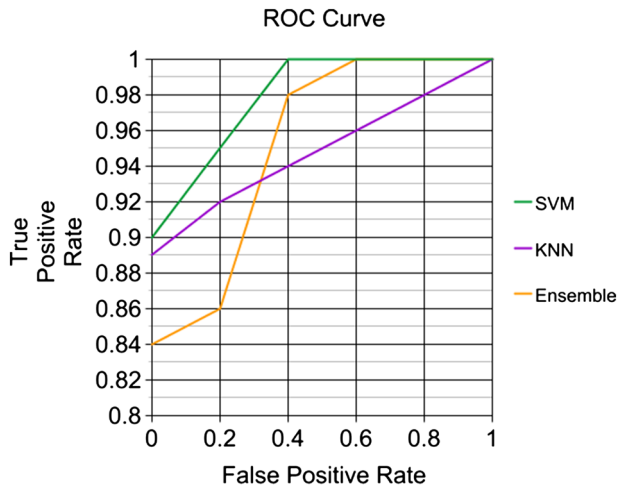


Fig. 6 The ROC curves of all three experimentally evaluated classifiers

Table 7 AUC values of the three experimentally evaluated classifiers

Classifier	AUC
SVM	0.98
k-NN	0.95
Ensemble	0.96

Table 8 Results of the comparison of our work and others' works

Work done/references	Domain	Size of data set	Type of learning	No. of features	No. of fish family	Classifiers used	Accuracy (%)
This work	Bangladesh	180	CML	14	6	SVM	94.2
						k-NN	92.6
						Ensemble	88
Fish species recognition using computer vision and a neural network [10]	NA	251	Neural network	NA	6	NA	95
Fish species classification by color, texture and multi-class support vector machine using computer vision [12]	China	540	CML	16	6	VBMSVM	97.96
						DAGMSVM	97.77
						LIBSVM	95.92
Fish recognition based on robust features extraction from size and shape measurements using neural network [13]	Malaysia	350	Neural network	18	20	NA	86

In order to evaluate the performances of three classifiers, a confusion matrix is formed for each classifier. The three confusion matrices of the three classifiers are displayed in Tables 3, 4 and 5. Accuracy, sensitivity, precision, FPR and FNR—these prominent performance metrics are calculated for each classifier, which is shown in Table 6. It can be seen from Table 6 that accuracy of SVM is 94.2%, which shows better accuracy than other two classifiers. Not only accuracy but also sensitivity, precision, FPR and FNR of SVM classifier are better than k-NN and Ensemble classifiers

We also use receiver operating characteristic (ROC) curves for making our claims more rigorous [23]. We use ROC curves for comparing the performance of classifiers. The area under the ROC curve (AUC_{ROC}) indicates comparison of classifier which is better on average. If the model is consummate, then AUC_{ROC} value must be equal to 1. The classifier which have a larger A_{ROC} , is better than other classifiers [23]. In this work, ROC curve of all three classifiers show in Fig. 6 and the area under curves are shown in Table 7. From Fig. 6, we can estimate that SVM performance is better than the other two classifiers. This statement becomes stronger by Table 7 in terms of AUC (Area Under The Curve) values.

6 Comparative analysis of results

For evaluating the quality of our proposed system in recognizing local fish, comparison has needed with some recently revealed incidental research outcome. In addition, it becomes challenging to have a right comparison of quality of different approaches because of the lack of using common database of sample of fish. There have been some research about fish recognition for past few

years but methodical and comparative performance evaluation based on practical assumption is not sufficient. Though there are some limitation, we have tried to review numerical results related to fish recognition to assess comparative quality to this work. Table 8 shows an overview of all methods of different works including this work.

In paper [10], height and width of fish at various location had been used as features and this measured value had been used in neural network. Network trained to recognized data and the accuracy of testing data was 95%.

In paper [12], for fish recognition, they used color and texture as features and two multi-class support vector machine used for classification. VBMSVM (Voting based Multi-Class Vector Machine) gave the best accuracy in their work which is 97.96%.

In paper [13], features were extracted by measuring the distance and geometrical measurement. Neural Network associated with black propagation algorithm was used for recognizing fish and overall accuracy was 86%.

Corresponding discussed scenario from the beginning of the section, obtained accuracy of this work is more than 94% for SVM Classifiers which turns out to be good and promising enough. In the starting of this section, we have mentioned that due to lack of similarity in image data set with different resolution, performance evaluation and the nature of intended application, it is not sensible to explicitly compare quality of our approach with other works.

7 Conclusion and future work

Fish recognition system based on machine vision has been represented in this paper. This experiment has been done with six species of fish and fourteen features. We have used PCA for feature selection. Three classifiers have been deployed in order to make a good recognition of fish. Among the classifiers, SVMs have produced good result achieving 94.2% accuracy.

As deep learning achieve nearly human-level performance on such type of computer vision tasks. In future we will intend to use deep learning algorithms. In this work, without any detection, using high resolution and some particular angles, pictures will be recognized. But in future, we want to do more work so that low resolution and any kind of angle pictures will be also identified. For the future work, it is planning to deal with massive data set of images covering more species of both freshwater and saltwater fishes of Bangladesh. There exists a promising work for future with a massive set of fish image data to cover a wider range of countries including Bangladesh.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Shamsuzzaman MM, MahmudullIslam M, Tania NJ, Abdullah Al-Mamun M, Barman PP, Xu X (2017) Fisheries resources of Bangladesh: present status and future direction. *Aquac Fish* 2(4):145–156
2. Arivazhagan S, Newlin Shebiah R, Nidhyandhan SS, Ganesan L (2010) Fruit recognition using color and texture features. *J Emerg Trends Comput Inf Sci* 1:90–94
3. Semary NA, Tharwat A, Elhariri E, Hassanien AE (2014) Fruit-based tomato grading system using features fusion and support vector machine. In: *IEEE conference intelligent systems*, pp 401–410
4. Lamon FG, Cervantes J, Ruiz S, Chau AL (11 August 2015) Color characterization comparison for machine vision-based fruit recognition. In: *International conference on intelligent computing*, pp 258–270
5. Zhang Y, Wang S, Ji G, Phillips P (2014) Fruit classification using computer vision and feed forward neural network. *J Food Eng* 143:167–177
6. Zhang B, Huang W, Gong L, Li J, Zhao C, Liu C, luang D (2014) Computer vision detection of defective apples using automatic lightness correction and weighted RVM classifier. *J Food Eng* 143:167–177
7. Sluzek A (2014) Machine vision applications to aquatic foods: a review. *J Food Eng* 11:167–177
8. Gümüş B, Balaban MÖ, Ünlüsayın M (2016) Machine vision in food recognition: attempts to enhance CBVIR tools. In: *Position papers of the federated conference on computer science and information systems*, vol 9, pp 57–61
9. Yanai K, Kawano Y (29 July, 2015) Food image recognition using deep convolutional network with pre-training and fine-tuning. In: *IEEE international conference on multimedia & expo workshops*
10. Storbeck F, Daan B (2001) Fish species recognition using computer vision and a neural network. *Fish Res* 51(1):11–15
11. Dowlati M, Ia M, Majid G, Seyed D, Mohtasebi S (2012) Application of machine-vision techniques to fish-quality assessment. *Trends Anal Chem* 40:168–179
12. Hu J, Li D, Duan Q, Han Y, Chen G, Si X (2012) Fish species classification by color, texture and multi-class support vector machine using computer vision. *Comput Electron Agric* 88:133–140
13. Alsmadi MK, Omar KB, Noah SA, Almarashdeh I (2010) Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements Using Neural Network. *J Comput Sci* 6:1059–1065
14. Tan P-N, Steinbach M, Kumar V (2006) *Introduction to data mining*. Addison-Wesley, Boston
15. Phillips D (2000) *Image processing in C*, 2nd edn. R & D Publications, Lawrence
16. RGB to HSV conversion: <https://www.pantechsolutions.net/image-processing-projects/matlab-code-for-image-retrieval>. Accessed 24 Jan 2019
17. Hughes GF (1968) On the mean accuracy of statistical pattern recognition. *IEEE Trans Inf Theory* IT-14(1):55–63
18. Keys R (1981) Cubic convolution interpolation for digital image processing. *IEEE Trans Acoust Speech Signal Process* 29(6):1153–1160

19. Fourier transform: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/fourier.html>. Accessed 24 Jan 2019
20. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans Syst Man Cybern SMC*–3(6):610–621
21. Malhi A, Gao RX (2004) PCA-based feature selection scheme for machine defect classification. *IEEE Trans Instrum Meas* 53(6):1517–1525
22. Habib MT, Rokonzaman M (2012) A set of geometric features for neural network-based textile defect classification. *Int Sch Res Netw* 2012:16
23. Habib MT, Majumder A, Jakaria AZM, Akter M, Uddin MS, Ahmed F (2018) Machine vision based papaya disease recognition. *J King Saud Univ-Comput Inf Sci*. <https://doi.org/10.1016/j.jksuci.2018.06.006>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.