Inteligencia Artificial 25(70) (2022), 13-32
doi: 10.4114/intartif.vol25iss70pp13-32

INTELIGENCIA ARTIFICIAL

http://journal.iberamia.org/

A Machine Vision Approach for Recognizing Coastal Fish

Afiq Raihan[1,A,*], Israt Sharmin[1,B], B M Marjan Khan[1,C], Md. Ismail Jabiullah[1,D], Md. Tarek Habib[1,E]

[1]Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh.

[A]afiq.cse@gmail.com, [B]ima.sharmin23@gmail.com, [C]marjan3085@diu.edu.bd, [D]drismail.cse@diu.edu.bd, [E]md.tarekhabib@yahoo.com.

* Corresponding Author: Afiq Raihan

Abstract Coastal fish is one of the prominent marine resources, which takes a necessary role in the economic growth of a country. Because of environmental issues along with other reasons, not only most of the marine resources are diminishing but also many coastal fishes are getting extinct gradually. As a result, the young peoples have insufficient knowledge of coastal fish. This issue can be solved with the use of vision-based technologies. To deal with this situation, a coastal fish recognition system based on machine vision is conceived, which can be approached by the images of coastal fish that are captured with a portable device and identify the fish to recognize fish. Numerous experimental analyses are executed to exhibit the benefit of this proposed expert system. In the beginning, the conversion of a color image into a gray-scale image occurs and the gray-scale histogram is developed. Using the histogram-based method, image segmentation is conducted. After that, a set of sixteen features comprising four classes is extracted to be fed to a classifier. For reducing the number of features, PCA is applied. To recognize coastal fish, five classical machine learning classifiers are performed, where *k*-NN provides a potential accuracy of up to 98.89%.

Keywords: Fish species recognition, Machine vision, Feature extraction, Principal component analysis, *k*-nearest neighbor, Performance metric.

1 Introduction

Fish is an ecumenical food in the glove containing a large amount of protein, calcium and phosphorus, and other mineral sources. Most Bangladeshis have to take fish to fulfill the need for protein in their daily life. Even so, fish products make up nearly 20% of all human diet foods consumed worldwide [1]. Geographically, Bangladesh has been considered the largest wetland of rivers and sea areas for fisheries after China and India in Asia [2], with a large suitable environment of the Bay of Bengal bordering India and Myanmar. According to the fisheries statistics report of Bangladesh [3], the gross domestic product (GDP) rate is 3.57%, which is 26.50% of agricultural GDP in the 2020-21 fiscal year, contributing a crucial role to the growth of the national economy. Despite the vast affluence of marine resources in Bangladesh, only 14.90% of total production comes from this sector [3]. There are over 32000 fish species globally, with approximately 401 fish species living on Bangladesh's sea land. [4]. The living place of fish is becoming a challenge to climate change, water pollution, and other human causes. Many of the fish species are reduced by the threat to the fish's ecosystem. Some of them are catfish ("Balitora" in Bangla), Ek Thota, Koi Bandhi, Darkina, Nalua Chandra, Lal Chandra, Catfish, Tilla, Shapla Pata, Red Salmon, etc. In this scenario, the future generation has no idea about the imperiled fishes.

ISSN: 1137-3601 (print), 1988-3064 (on-line)

In our study, an experiment is carried out on local coastal fish to recognize them using a machine vision-based technique. An expert system has been represented in this paper that not only processes coastal fish images but also recognizes the fish from the image. We have proposed a feature set so that coastal fish can be recognized. The features are extracted after segmenting the image from the original image. For the reduction of feature dimensionality, the principal component analysis (PCA) is used so that a meaningful comparison can be done. After that, five classical machine learning classifier techniques including *k*-nearest neighbor (*k*-NN), support vector machines (SVM's), linear discriminant, Naïve Bayes, and bagged trees are utilized.

In summary, the primary contributions of our research are:

- The main target is to identify the abolished fish using computer vision techniques, an off-the-shelf smart system.
- Basically, a machine learning problem is solved whereas the database, features, and methods are in properly well structured.
- Our proposed approach has performed a comprehensive outcome on our dataset to accomplish the goal.
- A progressive dataset has been built for possible improvement.

The remaining part of the following paper is categorized as: Section 2 demonstrates the present state of the research. The architecture of the suggested coastal fish recognizing expert system illustrates in Section 3. Section 4 contains the research methodology of our research. The data processing, feature extraction, and feature selection procedures have been described in section 5. Section 6 analyzes the outcomes and discussions in detail. Section 7 shows the comparative analysis of relevant works. Finally, Section 8 contains the conclusion, limitations, and future scope of this research work.

2 Literature review

Automated machine vision based object recognition is a common fact nowadays. Many research efforts have been made to identify the local fruit disease recognition, such as papaya [5], jackfruit [6], guava [7], and so on. On the other hand, there are few works on fish recognition among them. However, it is great to note that promising work has been done on local fish recognition in our country. In paper [8], the authors demonstrated a local fish recognition system on six different fish species. They used fourteen features, and the PCA algorithm was used to decrease the feature dimension. The outcomes were based on *k*-NN, SVM, linear discriminant, Naïve Bayes, and ensemble classifiers, and SVM classifier accuracy is 94.2% which is the highest accuracy. They employed a total of 180 photos for six species, dividing the dataset into 96 for training and 84 for testing. In this case, they need to expand the dataset to improve identification. Apart from that, the authors of article [2] showed the potentiality of the fishery sector and provided the future direction, scopes, guidance, and prospective scenario of Bangladesh.

In the paper [9], the author presented a fish recognition system for identifying invasive fish species. They demonstrated the evolution-constructed (ECO) features and the AdaBoost classifier. The system achieved 98.9% accuracy when 1049 photos of 8 different fish species were used. The researchers presented an integrated classification system termed fitness-scaled chaotic artificial bee colony (FSCABC) and feedforward neural network (FNN) in their research work [10]. First, they scaled the collected photos to 256×256 pixels and used the color histogram, texture, and shape attributes of fruits for 18 categories. Second, the reduction of the dimensions had been done using PCA. However, FSCABC–FNN system achieved 89.1% accuracy in fruit classification. The work [11] represented a multi-class support vector machine (MSVM) classifier for identifying species of fish where color and texture features are extracted. The image was acquired in 1024×768 pixels and an automated cropping system had been applied to prune the specified skin region from the main pictures in 512×512 pixels, while they utilized the smartphone due to its comprehensive usage. They obtained the features from the RGB, HSB color characteristics, and also texture features including grayscale histogram, gray level co-occurrence matrices (GLCM's), and wavelet transform. The authors constructed the one-against-one based multiclass support vector machines such as Voting Based Multi-class Support Vector Machine (VBMSVM), Directed Acyclic Graph Multi-class Support Vector Machine (DAGMSVM) and analyzed the predictive results with the LIBSVM software. The author of the paper [12] described that they merged the features based on size and form measurements, bearing in mind distance and geometrical tools. They concentrated on the recognition of landmark points and gave eighteen features to identify the fishes. The landmark point detection was used to determine the area and distance of fish from one location to another. They trained the neural network on 350 photos (257 for training and 97 for testing) and achieved an accuracy of 86%. They should enhance the big volume of data for greater neural network accuracy.

In [13], they claimed that the experiment was based on form and texture features and that the SVM classifier achieved an accuracy of 78.59% when compared to others such as *k*-NN, artificial neural network (ANN), and *K*-means clustering classifier. The accuracy was not adequate due to the dataset's limitation (150 samples). In the research paper [14], the authors presented an iterative combination-based feature selection strategy that optimizes classifier performance. For feature selection, two classifiers, Naive Bayes and Attributed Selected, were employed, and three classifiers, J48 decision tree-inducing method, multilayer perceptron (MLP), and SVM, were used to identify the species based on the characteristics. Efficient features such as geometric features, a bag of visual word models, and texture features were demonstrated in [15]. They obtained better accuracy on the conjugation of geometric features, a bag of visual word models. Recognizing fish species is accomplished using an ensemble learning method, i.e. the random forest.

From the above review, we can demonstrate that the coastal fish can bring nobility to the fishery of Bangladesh. But still, research work in this area has been unpleasant. In addition, most of the works are related to the recognition of fruits or fruit diseases. Even though there are not enough resources for better measurement and the relevant works for it.

3 Architecture of the expert system

Expert system architecture using machine vision to recognize local coastal fish is illustrated in Fig. 1. We concentrate on the mobile phone-based expert system for user usability and accessibility because the smartphone is the most extensively used electronic device. The expert system displays the hypothesis in which the user collects the image of the fish using a smartphone or a handheld device. Meanwhile, the user must download and install the Android app. The taken image is then supplied to our proposed expert system through the internet, and the receiver delivers the image to the data server for analysis, where we obtain the projected result. Finally, the user's smartphone or portable device will get the projected outcome.

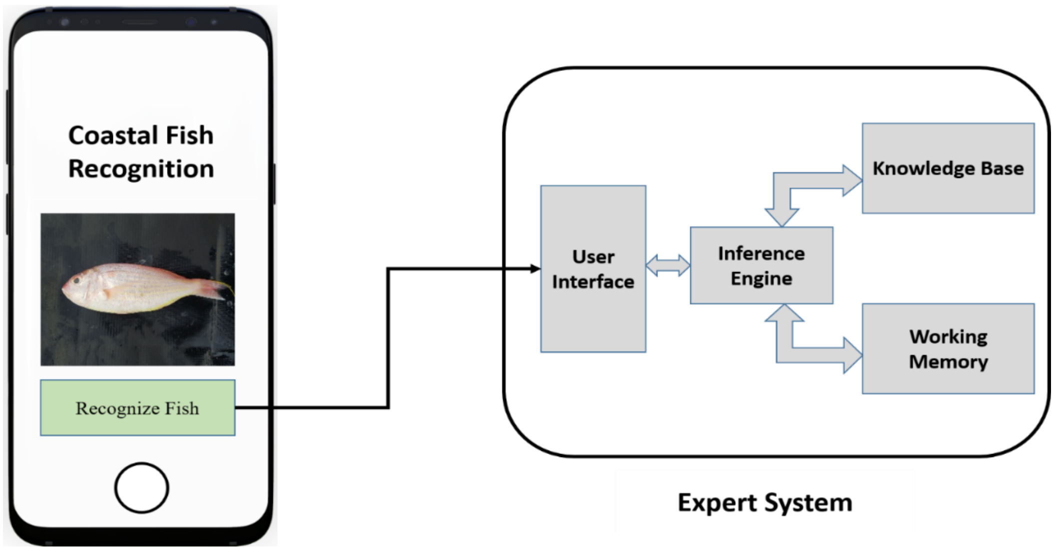

Fig. 1. The architecture of an expert system using machine vision for coastal fish recognition

4 Research methodology

Our approach for coastal fish species recognition using computer vision, as illustrated in Fig. 2, commences with the color images of coastal fishes. Since the images have been deadly captured from the local market, the research is restricted to the fixed orientation number. In the work, classical machine learning has been applied. It is inferred that the ideal classifier model is determined which produces the lowest generalized error [16, 17]. The classifiers that generate the optimal model error have been used. To learn more thoroughly by deep learning, the images of the data will be needed in an enormous quantity. But the huge number of data cannot possibly accomplish in this research. That's why traditional machine learning has been applied for fish recognition.

Initially, the fish image is restored as a fixed image in a 512×512-pixel image. Afterward, it is needed to transform the color image into a gray-level image. Considering the pixel values of red, green, and blue are symbolized as R, G, B of the original image, and w represents the grayscale's value for that exact point, under [17], w can be constructed as below:

$$w = 0.33 \times R + 0.56 \times G + 0.11 \times B \tag{1}$$

Now a gray-scale image is needed to segment by a histogram-based approach, such as the histogram peak technique [18]. The approach is most popular and simple for the segmentation of the image. Two threshold values are counted for converting it to a binary image which is symbolized as θ_L and θ_H. Then the values are optimized. This optimization is based on the images, and the values can be estimated by calculating each after separating the image section. Each pixel of the image is represented by $p(g, h)$ and the image is transformed into a black and white image in which each pixel is represented by $bi(p, q)$, as

$$bi(p,q) = \begin{cases} 1, & if\ \theta_L \leq pi(p,q) \leq \theta_H \\ 0, & otherwise \end{cases} \tag{2}$$

The binary images have two elements: object and background. Then the features have been extracted from the segmented image. After that, the object has been resurfaced to conduct color intensity features. As it is beneficial, color intensity has been also used for feature extraction. The RGB mean value has been measured from the color segmented image. After that, the RGB image is transformed into HSV to measure the mean value of HSV. It is known to all that HSV is more consistent for color analysis in feature extraction than RGB. If Hue is represented by H and specifies the character of the visual sensation that refers to the perception of color correlated with the influential colors, the value varies from 0 to 360°. Saturation is again symbolized by S and measures the degree of intensity or cleanness in which purification specifies how much white is applied to the 0 to 1 color range. Here, the brightness of a color is measured by V. So, this can be demonstrated from [19] in the following way:

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

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

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

Then, the other features are carried out from the segmented image. Here, sixteen features are combined in this research. The sixteen features are too much for the seven types of fish. To reduce the features, the feature selection is the optimal solution which is compared to the extracted features. The curse of dimensionality is a circumstance that is found during the analysis and organization of data in a high-dimensional space, yet this is not noticed as a third-dimensional physical space in a low-dimensional space [17]. First of all, this work is performed with sixteen features. These features are minimized to 11 features from 16 features because of the risk of dimensionality curse. To the fact of observation, PCA has been performed and the method has been completed because the primary objective of PCA is to carry out a new dimensional set that can identify quickly the data convertibility. When the model data set exceeds the 2 to 1000-item limit, the dimensional curse occurs frequently [20]. Since the research has been proposed for mobile-based applications, if the data has simplicity, less computational complexity, and cost reduction are present, this will be advantageous. In this scenario, the PCA algorithm has been applied to this work. Apart from this, due to the question of actual accuracy, it cannot be possible to reduce less than the eleven features. For this, feature selection is applied and minimized the features into eleven by using PCA.

Their features and selection of them are discussed elaborately in the next chapter. Fig. 2 portrays the layout of the expert system for recognizing fish species using machine vision. Such feature vectors are fed to SVM's, *k*-NN, linear discriminant, Naïve Bayes, and bagged trees – five prominent classical machine learning classifiers, which are obtained to perform. After that, the classifiers are deployed upon the set of trained data, and the predictive results of the classifiers can be measured on the test data set. The efficiency of the classifiers may not be examined from only the accuracy measurement. False-positive (*FP*), false-negative (*FN*), true-positive (*TP*), and true-negative (*TN*) values are recorded as a binary confusion matrix-like two-class problem [21]. For the confusion matrix of multiple classes, the matrix can be constructed into $n \times n$ ($n > 2$) dimension, bringing the total number of n rows, n columns, and $n \times n$ entries. As the demonstration of the multiclass matrix described in [21], the results of *FP*'s, *FN*'s, *TP*'s, and *TN*'s for class $i = 1, 2, 3, \ldots\ldots, n$ are represented as follows:

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

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

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

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

The actual confusion matrix holds the mean values of n confusion matrices for each category and is the dimension of 2×2. The accuracy, precision, sensitivity, false-positive rate (*FPR*), and false-negative rate (*FNR*) are computed in percentage utilizing the confusion matrix, as stated in [17], in the following way:

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

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

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

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

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

Therefore, the outputs of the five classifiers concerning these metrics are measured in this way and they can identify fish species. Then the receiver operating characteristic (ROC) curves are performed because of measuring the experimental performance of the classifiers [17]. In this perspective, the five traditional classifiers are performed for the assessment of throughput.

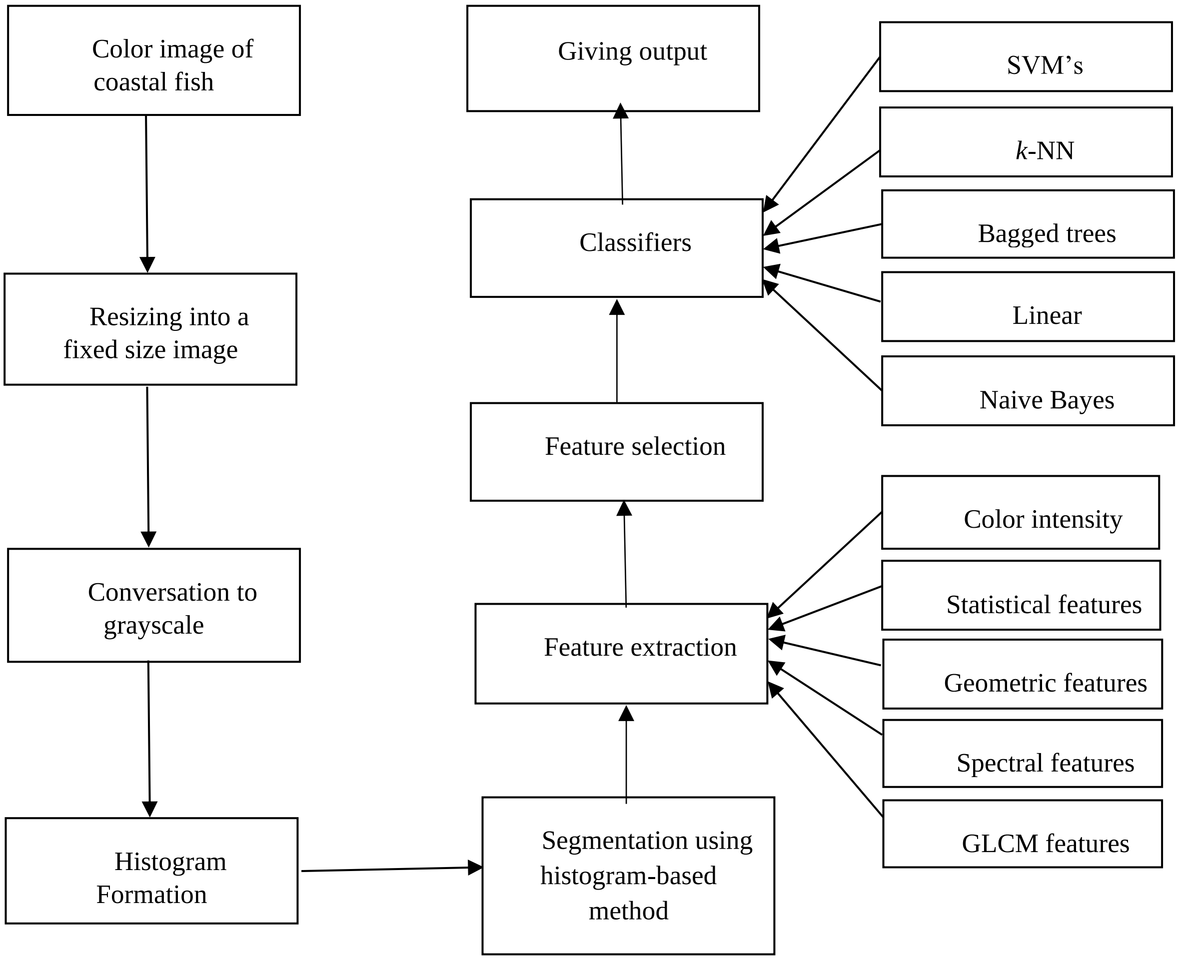

Fig. 2. Approach for recognition of local coastal fish.

5 Description of coastal fishes and features

The image data is collected from the southern part of Bangladesh. To analyze the performance, sixteen features are obtained from four categories, and selected eleven features after feature selection. All of the features and the feature selection approaches are presented broadly below in this section.

5.1 Description of coastal fishes

The diversification of fisheries is a global phenomenon. The fishes vary from one species to another. This diversification helps us to select some features to recognize the coastal fish. Here, seven types of coastal fish are used from all coastal saltwater fish. The local name of those fishes is Poa (Garra annandalei), Baila (Awaous guamensis), Loitta (Harpadon nehereus), Ilish (Tenualosa ilisha), Rupchandra (Pampus chinensis), Tailla (Eleutheronema tetradactylum), Koral (Lates calcarifer). The fishes are shown in Fig. 3. All of these species are sometimes accessible on the fish market in Bangladesh, which is legal. The fishes vary from one category to another with the shape, color, size, and feature. As a consequence, the features are combined depending on the various features of fish.

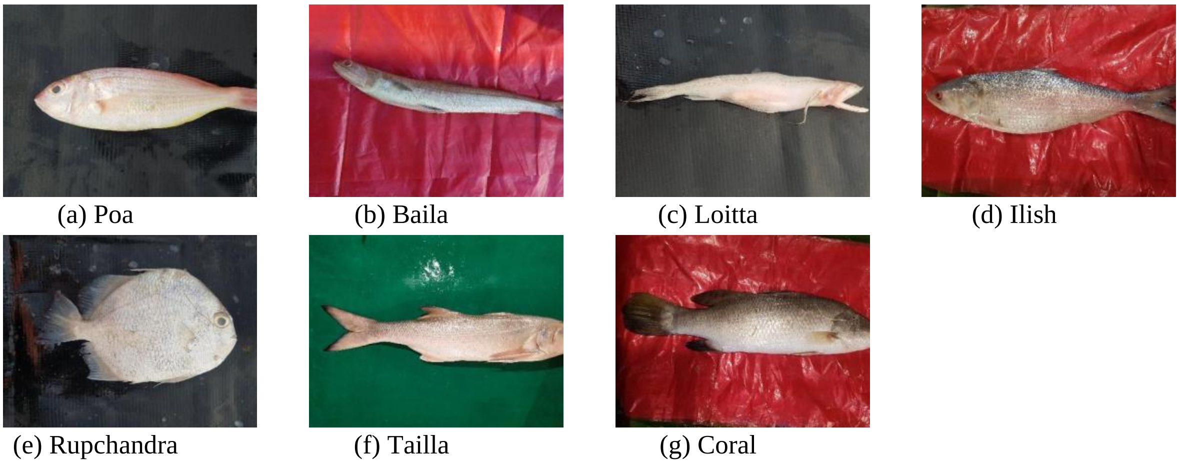

Fig. 3. Coastal fish species that are found in Bangladesh: (a) Poa (*Garra annandalei*), (b) Baila (*Awaous guamensis*), (c) Loitta (*Harpadon nehereus*), (d) Ilish (*Tenualosa ilisha*), (e) Rupchandra (*Pampus chinensis*), (f) Tailla (*Eleutheronema tetradactylum*), (g) Coral (*Lates calcarifer*).

5.2 Description of the image dataset

The research work, in this paper, on local coastal fish has been carried out as an experimental analysis by the approach of machine vision, for which the image data of fishes are collected from the local markets in the coastal area of Bangladesh. This is mentioned previously that the fishes are seen several times which is not forbidden by the authority of Bangladesh. The primary source of data collection is to capture images of fish which are taken by using the smartphone in different resolutions and angles. The seven species of coastal fish images are shown in Fig. 4 at various angles like 0°, 180°, 270° angles, and lower resolution. The total images of those fishes are 814 in this work. The class imbalance in the data set occurs in many systems in the real world when samples of all the classes are not distributed equally [17]. Since it become hard to find certain fish species at the time of this research work, the class became unbalanced. That's why the same number of images for all species of fish cannot be used in this research. Table 1 presents the total number of fish images separately by their species.

Table 1: Distribution of the fish image data

Fish	Frequency
Poa	150
Baila	120
Loitta	150
Ilish	120
Rupchandra	150
Tailla	48
Coral	76
Total	814

Fig. 4. The seven species of coastal fish images at various angles and lower resolution. (a) Name of Fish, (b) Original image, (c) Image of 180° angle, (d) Image of 270° angle, (e) Low-resolution Image.

5.3 Extracted features

For categorization, a set of sixteen (16) features from seven different fish species has been extracted. The feature set includes five categories of features such as color intensity, statistical, geometric, spectral, and GLCM [5], [8], [9], and [14]. These features are presented more elaborately below the section.

5.3.1 Color intensity features

When it comes to choosing features, the most important thing to keep in mind is the color model initially. The main image has been stored in the buffer. After that, the RGB and HSV mean values have been measured from the stored buffer image and the segmented images are not needed for those mean values. In this work, backgrounds have been utilized. The red, green, and blue components of color can be calculated using the RGB color model, which uses numerical representation. Every color has an independent value from zero to the highest value. We used the formula, as shown in (15), to determine the components of Red, Green, Blue, and the mean value of RGB Space (μ_{RGB}). If N_{FI} represents the number of pixels measured in a color image of fish (*FI*), according to [11], the mean value of RGB is:

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

Where $I(x, y)$ is the value of pixel of RGB at position (x, y). After that, the image is converted into HSV space as this is the most preferred approach for color analysis. The mean value of HSV space (μ_{HSV}) is the average of saturation, hue, and brightness in HSV space and is proportional to the number of pixels measured from the color image (*FI*), according to [11], the mean value of HSV is:

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

Where $I(x, y)$ denotes the value of a pixel of HSV at position (x, y). The mean of RGB and the mean of HSV are two measures of color intensity that are commonly used in feature selection.

5.3.2 Statistical features

Some statistical features are applied in our research and explained thoroughly here.

1. Mean: The object regions and the background without the object are identified by the *N* and *M* respectively. Whereas *G* symbolizes the representation of a pixel's gray-scale color.

$$\mu = \frac{\sum_{i=1}^{N} G_i}{N} \tag{17}$$

2. Standard Deviation: When object regions(N), gray-scale color intensity(G), and mean gray-scale color intensity(μ) are accounted for a pixel, the equation is as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(G - \mu)^2}{N} \tag{18}$$

3. Variance: When object regions(N), gray-scale color intensity(G), and mean gray-scale color intensity(μ) are accounted for a pixel, the equation is described as follows:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(G - \mu)^2}{N} \tag{19}$$

5.3.3 Geometric features

The five (5) features are computed for the geometric feature analysis after the statistical features analysis. Those features are mentioned below:

1. Height and width: A segmented image's height and width are referred to as the total maximum value of the segmented image's pixels of length and width.

2. Area: Area is one of the most significant and extensively utilized geometric elements for distinguishing an image. This function displays the precise area of an image as well as the pixels of all detectable regions that are visible. Fundamentally, it counts the total amount of pixels contained within the item. The area is designated by *A*, and $P[i, j]$ is a fish image. The equation, however, is as follows:

$$A = \sum_{i=1}^{m} \sum_{j=1}^{m} P[i, j] \tag{20}$$

3. Solidity: Solidity is another prominent geometrics feature of our study. The feature measures the density of an object. However, the structure of an object can be perceptible. The equation of solidity (S) is,

$$S = \frac{Nc}{2\pi r} \tag{21}$$

4. Convex: The convex feature is the pixel's number in the convex hull. It is perceived as the image's smaller shape.

5. Mean Intensity: The mean intensity means the value of pixels of an object which varies from 0 to 255. As the primary information is stored in pixels, pixel intensity is more crucial for image classification.

5.3.4 Spectral features

Fourier transform is used to sequentially calculate the complex data which is added for feature extraction. This is indicated in the description of complex analysis of different dimensions, fluctuations, and different phases of an object. The Discrete Fourier Transform (DFT) is a sort of Fourier Transform that consists of a collection of samples that represent the full spatial image. The frequency number organizes the pixel number in the spatial image. The two-dimensional DFT is discussed in [22],

$$F(m,n) = \sum_{x=0}^{L-1}\sum_{y=0}^{L-1} f(x,y)e^{-2\pi(\frac{mx}{L}+\frac{ny}{L})} \tag{22}$$

Where, $f(x, y)$ is referred to as the image in the domain of spatial, and each point value $F(m, n)$ is determined by calculating the image of spatial by combining the function and adding the output. Meanwhile, $F(0, 0)$ denotes the average brightness, whereas $F(L-1, L-1)$ represents the maximum frequency [22].

5.3.5 GLCM features

Numerous GLCM features, previously proposed by Haralick et al. [23], are also applied. Here, an image of two-dimension with $M \times N$ pixels and L gray levels is determined as $f(x, y)$. Two pixels in $f(x, y)$ is identified as (x_1, y_1) and (x_2, y_2), the angle between the two and the ordinate is θ, and d is the space between two pixels. So, the GLCM $P(i, j, d, \theta)$ denotes as [23]:

$$\text{Contrast:} \quad C = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2 P(i,j) \tag{24}$$

$$\text{Correlation:} \quad \rho = \frac{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} i.j.P(i,j) - \mu_x.\mu_y}{\sigma_x.\sigma_y} \tag{25}$$

$$\text{Energy:} \quad E = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P(i,j)^2 \tag{26}$$

$$\text{Entropy:} \quad S = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P(i,j)\log_2 P(i,j) \tag{27}$$

$$\text{Homogeneity:} \quad H = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{P(i,j)}{1+(i-j)^2}, \tag{28}$$

Where, the sums of mean and standard deviation are symbolized as μ_x, μ_y, σ_x, and σ_y for the row and column matrix entries of GLCM.

5.4 Feature selection

The purpose of feature selection is to identify the features and decrease the features [13] and [21]. PCA algorithm is used as a feature selection in this research work. We have sixteen features in this study that we minimize into eleven categories utilizing the PCA approach.

Generally, the PCA strategy interprets n vectors $(x_1, x_2, \ldots\ldots, x_i, \ldots\ldots, x_n)$ from d-dimensional space to n vectors $(x_1', x_2', \ldots x_i', \ldots., x_n')$, in an alternate d'-dimensional space as [24],

$$x_i' = \sum_{k=1}^{d'} a_{k,i}e_k \,, d' \leq d, \tag{29}$$

Where eigenvector is symbolized as e_k and $a_{k,i}$ are the projections of original vectors x_i [24].

6 Experimental evaluation

Fig. 5 shows the results of experimental observation for recognizing coastal fish using machine vision. First and foremost, the image of coastal fish must be captured. Following that, the acquired image is transformed into a 512×512- dimensional image. In the following segmentation, the fish is separated from the background, which was done using the histogram peak technique [25].

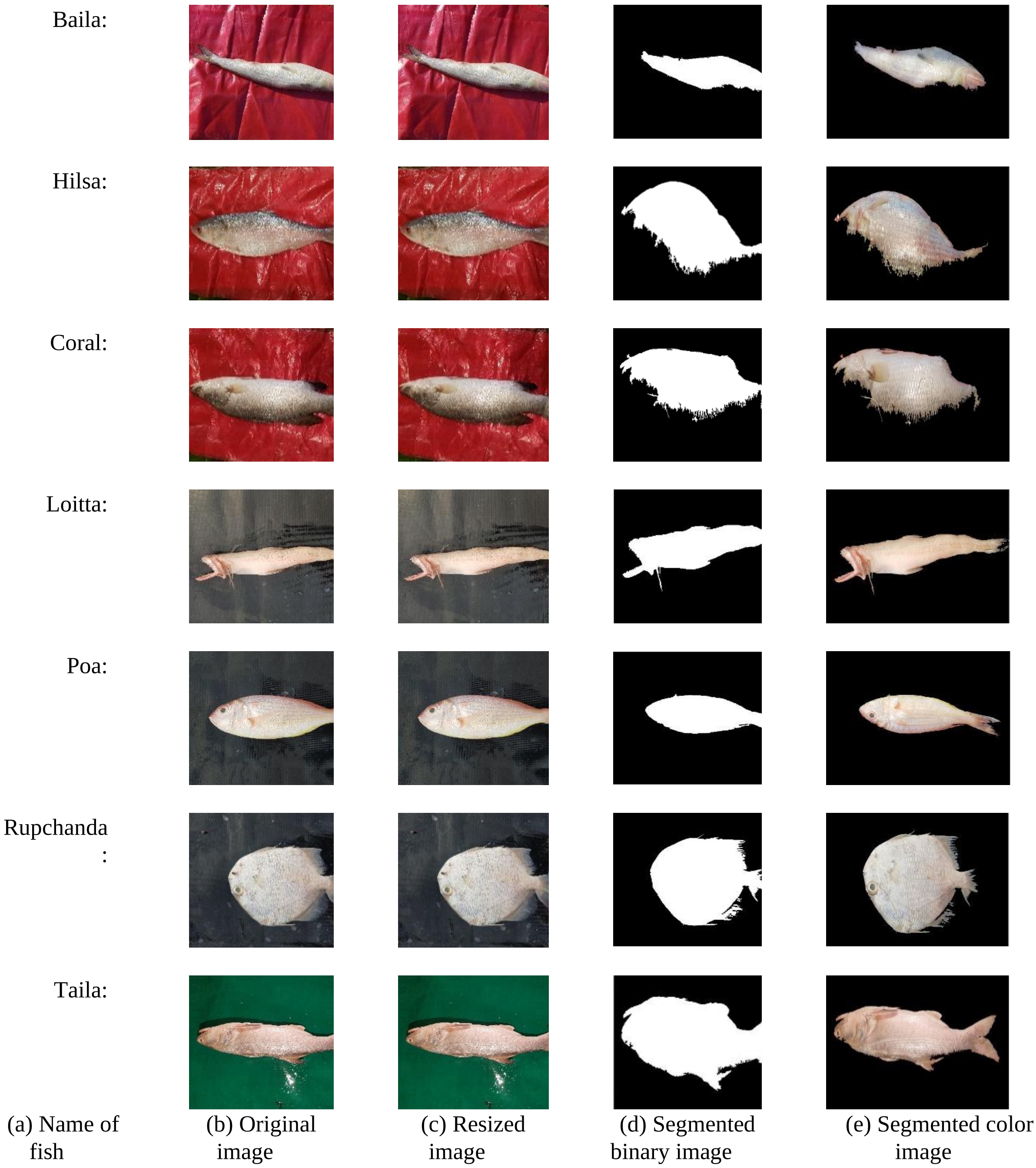

Fig. 5. Step-wise seven fish species image. (a) Name of fish, (b) Original image, (c) Resized image, (d) Segmented binary image (e) Segmented color image

Segmented fish images of seven classes have been shown in Fig. 5. Following the segmentation, sixteen features have been extracted. The features have been reduced into eleven features utilizing the PCA technique. In this way, the feature vector is established. Even though classifiers are also executed for sixteen features, noteworthy change has not been shown in this research work. To accomplish this work, eight hundred fourteen images of seven classes of coastal fish have been captured. In such a case, 814 becomes a considered data set, and the full set is split into two sets as training and testing data. The holdout method [17] was employed at random to eliminate the ratio of testing and training data. Here to mention that sample of data set is divided into five ratios of testing and training data sets which are 50% (407 images)-50% (407 images), 60% (489 images)-40% (325 images), 70% (570 images)-30% (244 images), 80% (652 images)-20% (162 images), 90% (732 images)-10% (82 images), for studying the comparison of accuracy in a different ratio. Implementation of validation sets is needed for avoiding model over-fitting problems, which implies having low generalization and low training error [25]. Under this procedure, two smaller subsets of training data have to be spliced from the actual set which is conducted for validation and training. Training data sets are used for classifier building, while testing data sets are applied to predict errors. The holdout approach was repeated three times to detect the trained classifier. Classifier performance is determined after each execution of the holdout technique where the test set has been applied. The computed average of thrice-found results is needed for building a multi-class confusion matrix.

In our study, five types of classifiers have been used such as SVM's, k-NN, linear discriminant, Naïve Bayes, and bagged trees. In this case, each parameter of classifiers is found with the optimized value which is specific. In Table 2. a list of elaborate parameter specifications of five classifiers is given.

Table 2: Explicit parameter specifications of five classifiers

Classifiers	Specification
SVM	Kernel: Linear
	C = 25007
k-NN	Metric of distance: Euclidean distance
	k = 1
Bagged trees	Each bag size (size of training data) = 100%
	Maximum depth of Trees = unlimited
	Randomly chosen attributes = 0
Linear Discriminant	The estimate of the class mean, $\hat{\mu}_k = \frac{\sum_{n=1}^{N} M_{nk} x_n}{\sum_{n=1}^{N} M_{nk}}$, where $M_{nk} \begin{cases} 1 \text{ if observation } n \text{ is from class } k \\ \text{otherwise } 0 \end{cases}$
Naïve Bayes	Distribution: Normal distribution = $f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
	Mean, $\mu_y = \frac{1}{N} \sum_i x^i$
	Variance, $\sigma_y = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{N}$

Following the sequence, the next step has been to analyze the performance of the classifiers. For each classifier, a confusion matrix is built so that the performance of the five classifiers can be evaluated. In this study, five classifiers have been applied both without (16 features) and with PCA (11 features) for comparing the accuracy of the result. Five ratios of data sets have been applied to study the accuracy in different data sets of five classifiers using without PCA and with PCA. In Fig. 6, the accuracy curve of the five classifiers without PCA illustrates that when the training data set is 70% (570 images) and testing data is 30% (244 images), the accuracy of the five classifiers is better than other ratios. In Fig. 7, the accuracy of five classifiers with PCA also shows that when the ratio of the train-test data set is 70%- 30%, the accuracy of recognizing coastal fish is higher than all other ratios.

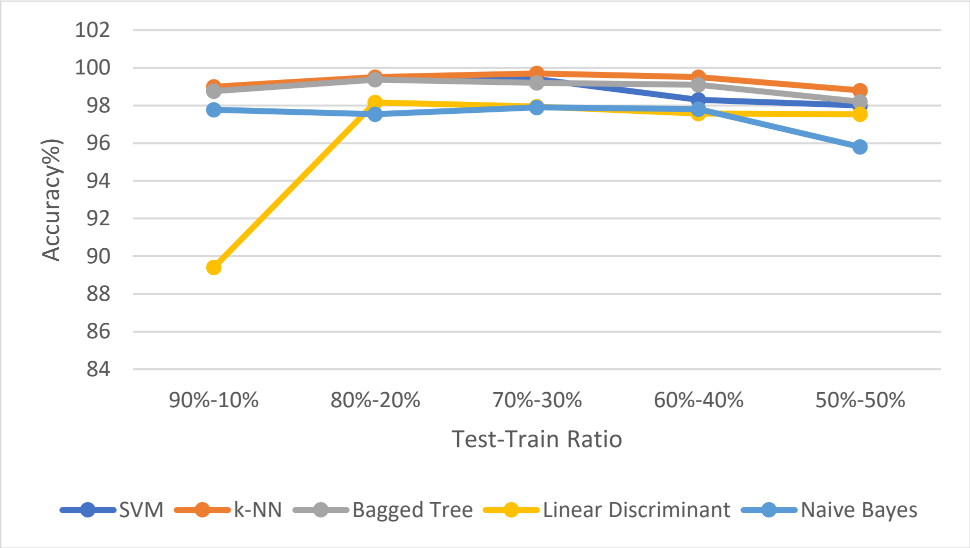

Fig. 6. The accuracy curve of five classifiers for five ratios without PCA.

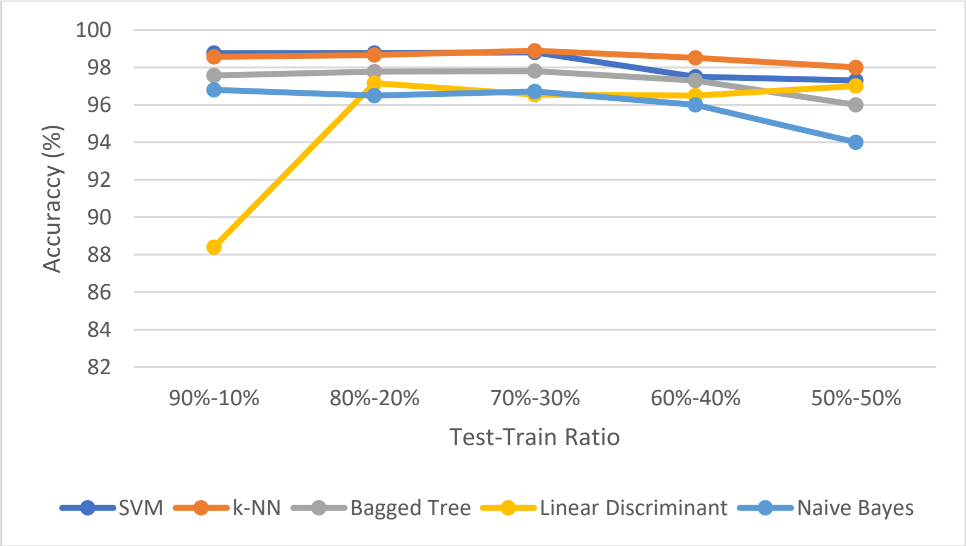

Fig. 7. The accuracy curve of five classifiers for five ratios using PCA.

The comparison between PCA and without PCA has also been done for five ratios of the dataset, which has been illustrated in Fig. 8. The graph shows that accuracy differs in terms of datasets and features and is up to 1%-2% higher when features have been 16. In the future, making mobile applications is a prime concern, so it will be better if the work has no complexity. For this motive, the work has been done using PCA.

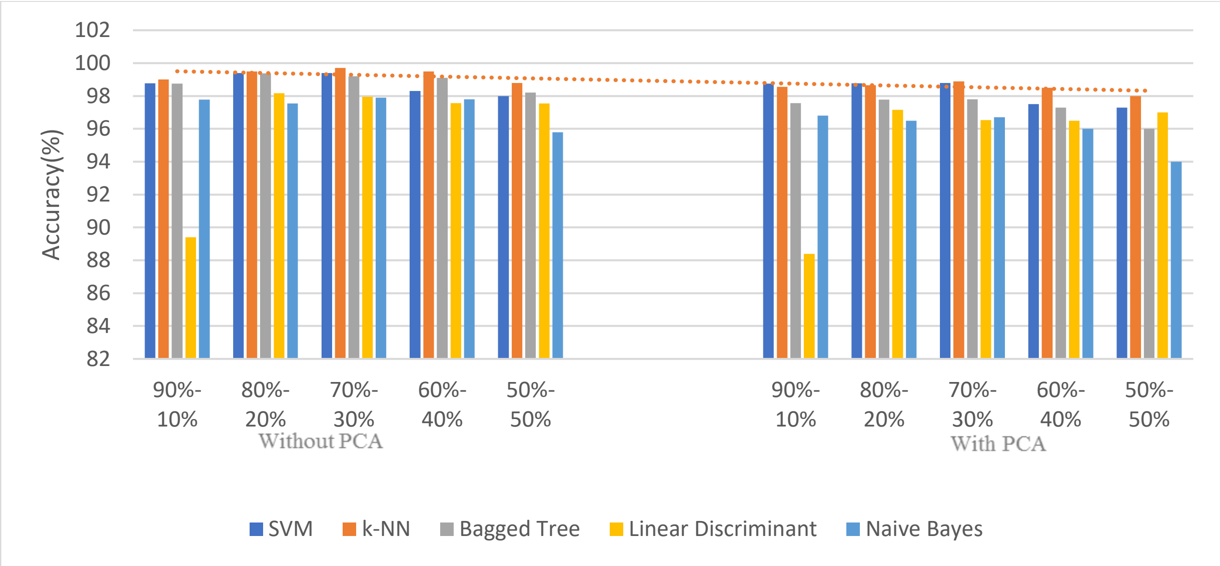

Fig. 8. Comparison graph of the accuracy of five classifiers for five ratios using without and with PCA

Table 3: The multiclass confusion matrix

Predicted/True Class	Baila	Hilsa	Koral	Loitta	Poa	Rupchanda	Taila
Baila	5	0.14	0	0	0	0	0
Hilsa	0	5.14	0	0	0	0	0
Koral	0	0	3	0.14	0	0	0
Loitta	0	0	0	6.43	0	0	0
Poa	0.14	0	0	0	6.29	0	0
Rupchanda	0	0	0	0	0	6.43	0
Taila	0	0	0	0	0	0	2

Most trials have indicated that *k*-NN classifier has obtained the optimal accuracy compared to other experimented classifiers. In this situation, we have shown most of the experimental data for the *k*-NN. When the training data ratio is 70% and the testing data ratio is 30%, the classifiers give the best result. As for this, the multiclass confusion matrix is formed and each class binary confusion matrices are also studied, which is illustrated in Table 3 and Table 4. In Table 5, the final binary confusion matrix is also created to evaluate the classifier's performance is demonstrated. The accuracy, sensitivity, specificity, precision, *FPR*, and *FNR* are also estimated for five classifiers which are shown in Table 6. From this table, we can see that the *k*-NN accuracy is

98.7% which is higher than other classifiers. It can also be noted that SVM and ensemble bagged trees have worse sensitivity, precision, *FPR*, and *FNR* than *k*-NN.

Table 4: Binary confusion matrices for each species

Class	Matrix				Class	Matrix				Class	Matrix			
Baila			Predicted Class		Hilsa			Predicted Class		Koral			Predicted Class	
			+	-				+	-				+	-
	Actual Class		5	0.14		Actual Class		5.14	0		Actual Class		3	0.14
			0.14	29.6				0.14	29.6				0	31.6

Class	Matrix				Class	Matrix				Class	Matrix			
Loitta			Predicted Class		Poa			Predicted Class		Rupchanda			Predicted Class	
			+	-				+	-				+	-
	Actual Class		6.43	0		Actual Class		6.29	0.14		Actual Class		6.43	0
			0.14	28.2				0	28.4				0	28.4

Class	Matrix			
Taila			Predicted Class	
			+	-
	Actual Class		2	0
			0	32.9

Table 5: The complete confusion matrix for all seven classes

		Predicted Class	
		+	-
Actual Class	+	4.89	0.06
	-	0.06	29.82

Table 6: Experimental results of the five classifiers

Classifier	Accuracy	Sensitivity	Precision	False Negative Rate	False Positive Rate	F1 Score
SVM	98%	97.50%	97.10%	2.50%	1%	98%
k-NN	98.70%	98%	97%	2%	1%	98.85%
Bagged tree	97.54%	97%	96.80%	3%	1.25%	97.85%
Linear Discriminant	96.20%	97.30%	97%	2.50%	1%	98.10%
Naive Bayes	96.71%	97.50%	97.10%	2.49%	1%	98.00%

To make this evaluation more rigorous, ROC curves have also been applied [5]. ROC curves are used to compare the performance of classifiers and this comparison can be perceived from the area under the ROC curve, AUC_{ROC}. The maximum and minimum values of AUC_{ROC} equals 1 and 0.5, respectively. The classifier with a high value of AUC_{ROC} means better performance than the others [5]. In Fig. 9, the ROC curves of all five classifiers are displayed and the corresponding AUC_{ROC} values have been shown in Table 7. We can notice that *k*-NN performance is preferable to the other two classifiers from Fig. 7. By Table 7 regarding AUC_{ROC} values, this statement becomes stronger.

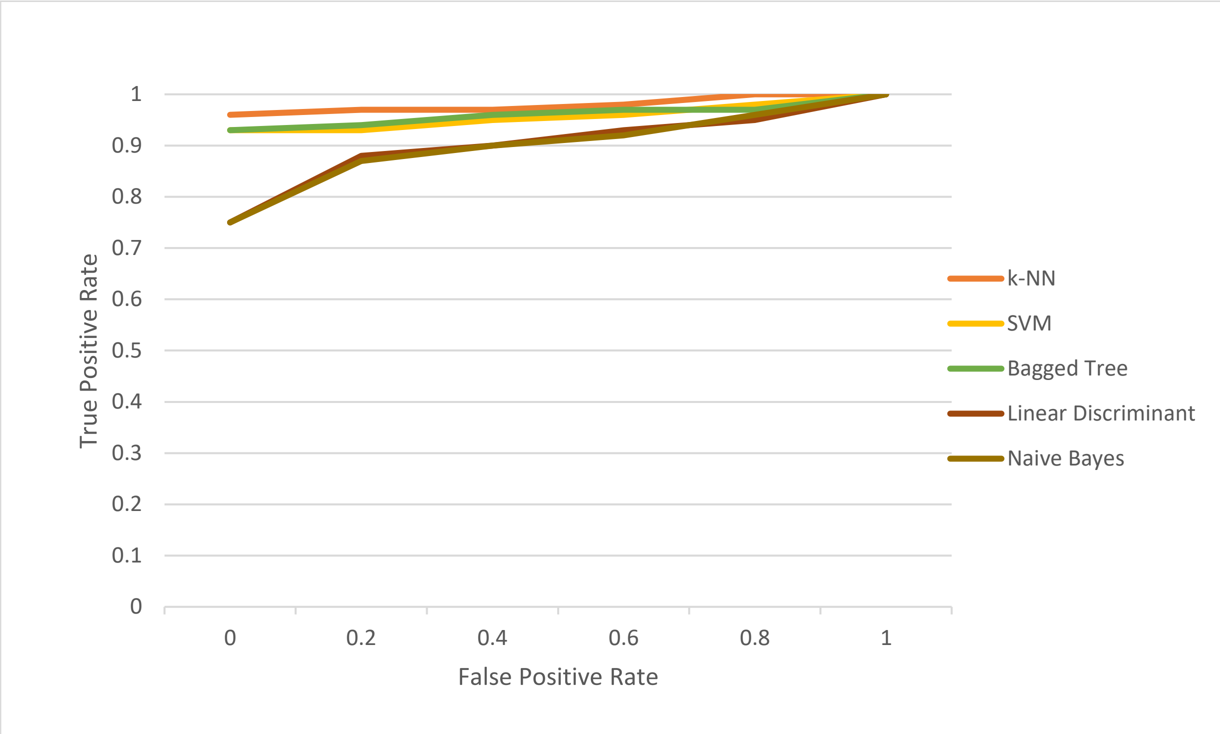

Fig. 9: The ROC curves for comparing five classifiers.

Table 7: AUC_{ROC} values comparing five classifiers

Classifier	AUC_{ROC}
SVM	0.97
k-NN	0.98
Bagged trees	0.96
Linear Discriminant	0.94
Naive Bayes	0.96

To accomplish the research work, the hardware requirements for the experiments are as follows:

- Processor: Intel core i3
- Ram: 8 GB
- System: 64-bit operating system
- Operating System: Windows 8.1

7 Comparative analysis of results

This research has been done with the machine vision approach for fish recognition on seven species. For the analysis of the proposed expert system, some recent works have been required for comparison with this study. Furthermore, the comparison of the performance evaluation on different approaches is difficult due to the absence of data equality and various image data in different resolutions. In recent years, there have been some reputed research works for fish recognition but the evaluation of the comparative and methodological performance on the practical analysis of the assumptions is not quite adequate. Although there are a few constraints, the analytical quantitative findings relevant to fish identification have been shown for evaluating the comparative output. The comparative performance of our findings is shown in the following Table 8 in detail.

Table 8: The performance of the comparative analysis of related works

Accomplished Work	Application Domain	Dataset Size	No. of Features	No. of Class	Classifier	Accuracy
This work	[1]CML	814	11 (with PCA)	7	SVM *k*-NN Bagged Tree Linear Discriminant Naïve Bayes	98.80% 98.89% 97.80% 96.54% 96.71%
			16 (without PCA)		SVM *k*-NN Bagged Tree Linear Discriminant Naïve Bayes	99.40% 99.70% 99.20% 97.95% 97.90%
Sharmin et al. [8]	CML	180	14	6	SVM *k*-NN Ensemble	94.2% 92.6% 88.00%
Hu et al. [11]	CML	540	16	6	VBMSVM DAGMSVM LIBSVM	97.96% 97.77% 95.92%
Zhang et al. [9]	CML	1049	[2]NA	8	AdaBoost	98.9%
Ogunlana et al. [13]	CML and Deep Learning	150	6	2	SVM *k*-NN ANN *K*-means	74.32% 52.69% 60.01% 50.97%
Hnin et al. [14]	CML and Neural network	1516	5	20	MLP(ANN) SVM J48	82.61% 99.13% 93.04%
Tan et al. [26]	CML and Deep Learning	174	NA	7	ANN SVM RF *k*-NN	91.14% 88.57% 89.83% 85.60%

[1]CML: Classical Machine Learning

[2]NA: Not Applicable

It can be noticed from Table 8 that the authors in the research work [8] have analyzed the fourteen features for the SVM, *k*-NN, and ensemble classifiers and found 94% accuracy for the SVM classifier. In paper [11], the authors have described a novel classifier approach in which VBMSVM obtained the best accuracy and DAGMSVM had the highest time efficiency.

The research work [9] has presented some ECO features and the AdaBoost classifier for invasive fish identification. The classification performance has been evaluated with an accuracy of 98.8%. The experiment in [13] was accomplished utilizing ANN, *k*-NN, SVM, and *K*-means clustering classifiers, with the SVM classifier accuracy being the greatest (78.59%), which was insufficient owing to the dataset's limitation. In paper [14], Iterative combination-based feature selection approaches such as naive Bayes and Attributed Selected were performed to extract the features, and the SVM classifier achieved 99.13% accuracy. In [26], they figured out the three traditional and deep characteristics, along with eight machine learning techniques. The ANN acquired 91.14% accuracy for deep features.

Concerning the discussed scenario at the beginning of this section, our acquired accuracy of the *k*-NN classifier is more than 98%, which is quite promising and optimistic. It is mentioned at the beginning of the section that the dataset of various works has the absence of data equality with the different resolutions. So, the evaluation of performance is varied in different works, and comparing our research work method with others is not rational.

8 Conclusion and future work

A systematic approach for coastal fish species recognition using machine vision has been proposed in this research paper. sixteen features are extracted from images of four categories of seven coastal fish species after segmenting the images using the histogram peak technique. The algorithm of PCA is also utilized to minimize the feature dimension. After that, three cutting-edge classifiers have been applied expecting good performances. The *k*-NN has delivered the best performance acquiring 98.89% accuracy, which is good and reckoning enough for fish recognition. Nowadays deep learning is being used in different research domains. So, the deep neural network is a prospective option in this regard considering the performance evaluation. A promising future work can be accomplished using deep learning with huge data including more categories of coastal fish and local freshwater fish species of Bangladesh.

Acknowledgments

The authors would like to convey our gratitude to those who help us to create our dataset.

References

[1] Cesarettin Alasalvar, Fereidoon Shahidi, Kazuo Miyashita, and Udaya Wanasundara. Handbook of Seafood Quality Safety and Health Applications. John Wiley & Sons, 2011.

[2] Md. Mostafa Shamsuzzaman, Mohammad Mahmudul Islam, Nusrat Jahan Tania, Md. Abdullah Al-Mamun, Partho Protim Barman, and Xiangmin Xu. Fisheries resources of Bangladesh: Present status and future direction. Aquaculture and Fisheries, *2*(4):145-156, 2017.

[3] Fisheries Resources Survey System (FRSS), Department of Fisheries. Yearbook of Fisheries Statistics of Bangladesh, 2020-21. Bangladesh: Ministry of Fisheries and Livestock, 38:138, 2022.

[4] Fish, Available online: http://en.banglapedia.org/index.php?title=Fish [Last accessed: March 20, 2022]

[5] Md. Tarek Habib, Anup Majumder, A. Z. M. Jakaria, Morium Akter, Mohammad Shorif Uddin, and Farruk Ahmed. Machine vision based papaya disease recognition. Journal of King Saud University-Computer and Information Sciences, 32(3):300-309, 2020.

[6] Md. Tarek Habib, Md. Jueal Mia, Mohammad Shorif Uddin, and Farruk Ahmed. An in-depth exploration of automated jackfruit disease recognition. Journal of King Saud University-Computer and Information Sciences, 34(4):1200-1209, 2020.

[7] Md. Rasel Howlader, Umme Habiba, Rahat Hossain Faisal, and Md. Mostafijur Rahman. Automatic recognition of guava leaf diseases using deep convolution neural network. In international conference on electrical, computer and communication engineering (ECCE), pages 1-5. IEEE, 2019.

[8] Israt Sharmin, Nuzhat Farzana Islam, Israt Jahan, Tasnem Ahmed Joye, Md. Riazur Rahman and Md. Tarek Habib. Machine vision based local fish recognition. SN Applied Sciences, 1(12):1-12, 2019.

[9] Dong Zhang, Dah-Jye Lee, Meng Zhang, Beau J. Tippetts, and Kirt D. Lillywhite. Object recognition algorithm for the automatic identification and removal of invasive fish. Biosystems Engineering, 145:65-75, 2016.

[10] Yudong Zhang, Shuihua Wang, Genlin Ji, and Preetha Phillips. Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering, 143:167-177, 2014.

[11] Jing Hu, Daoliang Li, Qingling Duan, Yueqi Han, Guifen Chen, and Xiuli Si. Fish species classification by color, texture and multi-class support vector machine using computer vision. Computers and electronics in agriculture, 88:133-140, 2012.

[12] Mutasem K. Alsmadi, Khairuddin B. Omar, Shahrul A. Noah, and Ibrahim Almarashdeh. Fish recognition based on robust features extraction from size and shape measurements using neural network. Journal of Computer Science, 6(10):1088, 2010.

[13] S. O. Ogunlana, O. Olabode, S. A. A. Oluwadare, and G. B. Iwasokun. Fish classification using support vector machine. African Journal of Computing & ICT, 8(2):75-82, 2015.

[14] Than Thida Hnin, and Khin Thidar Lynn. Fish classification based on robust features selection using machine learning techniques. In Genetic and Evolutionary Computing, pages 237-245. Springer, Cham, 2016.

[15] Takeshi Saitoh, Toshiki Shibata, and Tsubasa Miyazono. Feature points based fish image recognition. International Journal of Computer Information Systems and Industrial Management Applications, 8:12-22, 2016.

[16] Sasi Atia, and Khaled Shaalan. Increasing the accuracy of opinion mining in Arabic. In first international conference on arabic computational linguistics (ACLing), pages 106-113. IEEE, 2015.

[17] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to data mining. Pearson Education India, 2016.

[18] Dwayne Phillips. Image processing in C. Lawrence: R & D Publications, 1994.

[19] RGB to HSV conversion, https://www.pantechsolutions.net/image-processing-projects/matlab-code-for-image-retrieval. [Last accessed on March 21, 2022.

[20] Gordon F. Houghes. On the mean accuracy of statistical pattern recognition. IEEE Trans. Inform. Theory, 14(1):55-63, 1968.

[21] Robert Keys. Cubic convolution interpolation for digital image processing. IEEE transactions on acoustics, speech, and signal processing, 29(6):1153-1160, 1981.

[22] Fourier transform: https, Available online: https://homepages.inf.ed.ac.uk/rbf/HIPR2/fourier.htm [Last accessed March 20, 2022]

[23] Robert M. Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics, 6: 610-621, 1973.

[24] Arnaz Malhi, and Robert X. Gao. PCA-based feature selection scheme for machine defect classification. IEEE transactions on instrumentation and measurement, 53(6):1517-1525, 2004.

[25] Md. Tarek Habib, and M. Rokonuzzaman. A set of geometric features for neural network-based textile defect classification. International Scholarly Research Notices, 2012.

[26] Hui Yuan Tan, Zhi Yun Goh, Kar-Hoe Loh, Amy Yee-Hui Then, Hasmahzaiti Omar, and Siow-Wee Chang. Cephalopod species identification using integrated analysis of machine learning and deep learning approaches. PeerJ 9: e11825, 2021.