

RESEARCH

Open Access



Fish species classification using a collaborative technique of firefly algorithm and neural network

Pooja Prasenan* and Chethamangalathu Damodharaprabhu Suriyakala

*Correspondence:
poojaprasenan22@gmail.com

School of Ocean Engineering
and Underwater Technology
(SOEUT), Kerala University
of Fisheries and Ocean Studies
(KUFOS), Kochi, Kerala, India

Abstract

Monitoring various Fish Species and its distribution of the species obtains a primary significance in receiving the insights to marine ecological-system. After this, visual classification of those species would aid in tracing out the movement and yield the patterns and trends in fish activities, which provides in depth knowledge of the species. Unconstrained under-water images pose highly variations because of the fish orientation changes, Light-intensities, similarity in fish patterns and fish shapes. This would create the greater challenge for Image-processing techniques in accurate classification of Fish species or the Fish classes. Hence, for this reason, Underwater Image Enhancement is implemented in combination of Morphological-operations in pre-processing method. The pre-processed image is then subjected to feature extraction process by using Speed-up Robust Feature algorithm. This is followed by Firefly Algorithm, applied for optimization of Region of interest selection in the selected-features. For the categorization of Fish-species, PatternNet is a technique which is employed, in classifying 10,000 marine fish-images to five categories (*Dascyllus reticulatus*, *Plectroglyphidodon dickii*, *Chromis chrysura*, *Amphiprion clarkii*, and *Chaetodon lunulatus*). The Efficiency of the proposed-framework is performed in terms of Classification accuracy, execution time, precision value, F-measure and recall factors with respect to various categories of fish species. The comparison of the proposed-framework is also assessed with the other existing methods. 98% of accuracy rate in classification was produced by the evaluation results of the proposed framework with a lesser average computation time of 3.64 s upon different tested images. Thus, the higher efficiency of the proposed framework is proved by the outcomes of the study.

Keywords: Marine fish classification, Under-water image enhancement (UWIE) algorithm, Speed-up robust feature (SURF), Firefly algorithm (FFA), PatternNet

1 Introduction

Underwater Fish imagery in natural habitat faces several challenges such as image-distortions in noise and poor image-quality, variations in lightning, and background variations implying image nonlinearity in distribution of images. Hence these challenges ought to be addressed using enhanced nonlinear morphological operations for representing the complex feature-extraction process in Deep-learning approach. The

technique UWIE- is implemented in combination of Morphological-operations in pre-processing method. The classification of Underwater images can be performed by PatternNet is technique in classifying 10,000 marine fish-images to five categories. In this study, SURF-Speed-up Robust Feature algorithm is presented for extracting robust, non-variant, unique fish dependent features prior to classification [1]. The optimization of ROI-selection in the extraction features was applied through FFA. Conservation of marine life is inevitable in the present scenario as we are facing huge marine ecological scarcity which is caused due to overfishing, marine pollution, climate changes, etc. [2]. Also, we know that major percentage of population of Kerala depends on the marine life for their livelihood. Moreover, statistics shows that out of 80,500 species of living vertebrates nearly half is fish species. Besides this it is also known that 70.9% of earth's surface constitutes 97% of earth water, thus from the above facts we can assume that how important it is to protect and safeguard marine life [3]. The protection of fish species manually is a gruelling task and hence automated techniques are used. Real-time image capturing and video processing capabilities had significantly eased the screening of the most hazardous and least accessible areas such as underwater environment and sea-bed. Monitoring such environments is not only too difficult but is highly risky as well. The major task associated with the unpredictable lively environment of rivers, lakes and oceans is to capture high quality images to aid in further analysis. Moreover, marine life has a splendid variety of organisms that further challenges its apt recognition and categorization. In this context in order to minimize human efforts, numerous ways have been developed to capture and improve images of sea-bed, coral reefs, deep sea fishes and marine creatures [4].

For helping researchers incorporated in biomass estimation, underwater archaeology along with sea-bed analysis, automated detection and analytical methods which comprises machine learning and Deep Learning (DL) techniques proved to be very helpful in these situations. Recently, Awalludin et al. had presented a survey of available image processing techniques to de-noise, sharpen, de-blur and smoothen the underwater images [5]. Image processing has shown wider applications to analyze fish quality [6], abundance [7] in addition to length–weight measurement, detection [8], tracking [9], counting [7] and classification [8]. Additionally, Latest statistics shared by Food and Agriculture Organization of the United Nations states that total fish production will rise from 179 million tonnes to 204 million tonnes by 2030 in which Asiatic region will be the major contributor. Figure 1 summarizes the projected regional fish growth by state of world

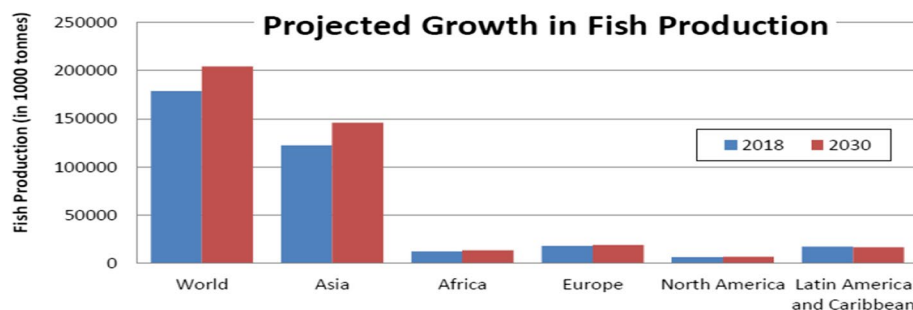


Fig. 1 The growth of fish production among different continents by 2030

fisheries and aquaculture 2020 [10] that reflect the growing need of precise classification of various types of fishes.

The tremendous growth in fisheries raises the need to monitor and manage fishes with techniques that could offer a safe and non-destructive sampling of fish species. Researchers had proposed a plenty of work in descriptor-based recognition of fishes and improving the image quality of underwater images using Scale Invariant Feature Transform (SIFT) with SURF [10], Multi-Objective Particle Swarm Optimization [11], content-based image retrieval [12], shape-based detection in underwater images [13] and combination of color and texture [14]. Some researchers also tried to categorize underwater fish images using Support Vector Machine (SVM) [10–15] and its variants [16]. But, poor visibility along with accessibility constraints of a lively marine environment limits the efficiency of these techniques. There are variations in the water turbidity, varying light luminosity, background confusion, along with the orientation of species and fish movement. In the present method, these challenges are discussed with the proposed UWIE algorithm along with the implementation of morphological operations and FFA-centered optimizations. The classification accuracy of the PatternNet classifier is further enhanced by SURF-centered feature extraction.

The authors have divided the paper into 5 sections including introduction. Section 2 summarizes the research pertaining to fish classification, Sect. 3 summarizes the proposed technique for fish classification and Sect. 4 discusses the achieved results. Section 5 concludes the paper and cited work is listed under references.

2 Similar research work for comparison

Plenty of research had been performed in the field of image-based fish classification in underwater environments. In this context, Fouad et al. [10] had proposed an automatic fish classification for Nile Tilapia fish. The major stress was given on feature extraction from the fish image using SIFT and SURF followed by SVM. Experimentation had shown that the fish classification achieved 94.4% detection accuracy using SURF and SVM. They also proposed to implement their classification approach to other fish species in future work [13]. Qin et al. [17] had proposed a framework based on Principal Component Analysis (PCA) in two layered architectures. In nonlinear layer, binary hashing and in feature extraction layer, block wise histograms were used. This was followed by spatial pyramid pooling to retrieve invariants in larger poses. In later stages, SVM was used as a machine learning classifier that demonstrated 98.64% classification accuracy against a real world fish recognition dataset. Despite the high accuracy, computational time and CPU consumption were still found challenging that needs to be addressed by the authors in their future work. Siddiqui et al. (2018) presented a Convolutional Neural Network (CNN)-based method for the classification of fine grained fish species using trained CNN for feature detection followed by classification performed using SVM. This strategy proved to exhibit 94.3% classification accuracy for the fishes captured off the Western Australia coast. Authors claimed that the automated classification proved to be cost-efficient in comparison to manual processing. Vilon et al. [18] proposed a CNN-based method to aid precise identification of various fish species using underwater fish images. The employed CNN was used against a test dataset of 1197 images representing 9 fish species and was trained over 900,000 fish images. Evaluation shows that different

performance is observed for individual fish species. Experimentation shows that *Dascyllus carneus* was only 4% correctly identified with one training dataset as compared to more than 90% when trained against three models. It was observed that for 18 fish species success rate ranged between 85 and 100%; however, for 3 fish species it was below 90% and for 9 species it was more than 95%. Overall, experimental analysis exhibited a true identification of 94.9% in comparison to manual fish identification with an average computation time requirement of 0.06 s as compared to manual identification of fishes. Lakshmi et al. [19] proposed an approach to detect and classify fishes using underwater data. Firstly, foreground detection was done using a Gaussian mixture model followed by bag-of-feature-based classification. Here, SURF-based features were clustered using k-means to develop a visual vocabulary. The testing was performed using Multiclass-SVM (MSVM) demonstrating 88.9% classification accuracy. Ahsan Jalal et al. [20] suggested a hybridized solution for combining optical flow along with Gaussian mixture models with YOLO deep neural network. It was a combined approach for detecting and classifying fish in unrestricted underwater videos. For capturing only the static along with evidently visible fish instances, YOLO-centered object detection systems were originally utilized. Using Gaussian mixture models and optical flow, temporal information was attained. On '2' underwater video datasets that was the LifeCLEF 2015 benchmark as of the Fish4Knowledge repository along with a dataset gathered by The University of Western Australia (UWA), the proposed system was assessed. Fish detection F-scores of 95.47% and 91.2% were attained by this method, whilst fish species classification accuracies of 91.64% along with 79.8% were obtained on both datasets correspondingly. More computational power was utilized by the proposed system since it comprised a complex machine learning tool when contrasted to traditional computer vision along with image processing approaches. Sebastien Villon et al. [21] suggested examining how DL limitations could be overcome by Few Shot Learning (FSL), which was a rising research field. FSL was centered on the principle of training a DL algorithm on 'how to learn a new classification problem with only a few images.' In this case study, the robustness of FSL was evaluated for discriminating '20' coral reef fish species with a range of training databases as of 1 image per class to 30 images per class, and contrasting FSL to a typical DL approach with thousands of images per class. This work found that the classic DL approach was outperformed by FSL in situations along with good classification accuracy was offered. The annotated images in this work were restricted.

Eko Prasetyo et al. [22] presented a Multi-Level Residual (MLR) as a new residual network strategy that combined low-level features of the initial block with high-level features of the last block by applying Depthwise Separable Convolution. Moreover, the MLR-VGGNet was implemented as a new CNN architecture inherited from VGGNet and strengthened by Asymmetric Convolution, MLR, Batch Normalization, and Residual features. Experimental results showed that MLR-VGGNet achieved an accuracy of 99.69%, outperformed original VGGNet relative up to 10.33% and other CNN models relative up to 5.24% on Fish-gres and Fish4-Knowledge dataset.

Muhammad Ather Iqbal et al. [23] presented an automated system for identification and classification of fish species. The developed approach was based on deep convolutional neural networks. It used a reduced version of AlexNet model comprises of four convolutional layers and two fully connected layers. The results showed that the model

with less number of layers had achieved the testing accuracy of 90.48% while the original AlexNet model achieved 86.65% over the untrained benchmark fish dataset. The inclusion of dropout layer had enhanced the overall performance of the model. Thus, the method has less memory consumption and it is also less computational complex. However, the method works well for smaller datasets.

Hafiz Tayyab Rauf et al. [24] developed a DL scheme centered on the CNN model for fish species recognition. To improve classification performance, VGGNet architecture was subjected to deep supervision by incorporating four CLs for the network's every level of training. On the Fish-Pak data set, an experimental comparison was conducted with various DL frameworks incorporating VGG-16 for transfer learning, three blocks VGG, two blocks VGG, one block VGG, AlexNet, LeNet-5, GoogleNet, and ResNet-50 to ensure the CNN architecture's superior working. The technique surpassed prevailing methods and attained state-of-the-art performance by a comprehensive empirical analysis. Since the performance was proved with fewer data, it was not reliable.

Ahsan Jalal et al. [25] suggested a hybrid resolution to merge Gaussian Mixture Models (GMM) and optical flow with YOLO deep neural network (NN), an incorporated scheme to perceive and categorize fish in free underwater videos. To capture stationary and apparent fish instances only, YOLO-based object detection methods were employed. Utilizing time-related information obtained via GMM along with optical flow; the system overcame YOLO's constraint, enabling to perceive freely swimming fish that were cloaked in the background. From the Fish4Knowledge repository, the LifeCLEF 2015 benchmark together with the dataset gathered by The University of Western Australia (UWA) was the '2' underwater video datasets in which the system is evaluated. Efficacy of the approach was procured from the outcomes of the model on a specified dataset. However, the system's computation time was lengthy.

3 Methodology

The proposed-framework elaborates the sequential steps evolved in Fish recognition and the classification techniques. This framework laid the baseline for biologists in analyzing the marine behavior of fishes and in assessing the under-water environment. In the initial stage of the framework, the input under-water fish images have been subjected to pre-processing techniques using UWIE. This is followed by other morphological-operations to handle the poor intensity and poor-light of those input under-water fish images. The Features of the Fish images have been extracted through SURF-Speed-up Robust Feature technique. The selected Features have been optimized in their intensity values through FFA. This firefly algorithm enhances the ROI selection process. And again, the extracted features though SURF-descriptors were optimized using FFA for attaining an optimal threshold value. These steps were all repeated for all the test-images and also carried out for the trained dataset of those fish images. The optimized features dataset of underwater fish images was trained by using classification algorithm PatternNet. This algorithm is implemented on the fish images for recognition and the classification of those fish images into the five predefined categories of fishes. The performance of the proposed-framework was assessed concerning classification accuracy, classification execution time, precision, recall values and F-measure values in those categories of classified fish types. The comparative analysis of the proposed-framework was performed

with that of the other existing classification methods. The efficiency of the framework has been estimated by this comparison. The steps involved in the methodology are described in Fig. 2.

Initially, the underwater fish images are preprocessed, and then the region of interest for the further processing are segmented out. From the segmented region, the most important features are extracted out, and then the extracted features are optimized thereby the dimensionalities of the features are reduced. These reduced features are given as input to the classifier, which efficiently classifies the fish types. Previous works and the methodology used in them are tabulated in (Table 1).

3.1 Fish image dataset

The Fish Recognition Ground Truth database consists of 27,370 fish image data that has been clustered into 23 categories with each category representing a unique fish species. The categorization was based on various features such as presence, number, position, shape of fins under the constructive guidance of marine biologists. The dataset is available online at <http://groups.inf.ed.ac.uk/f4k/GROUNDTRUTH/RECOG/> [20, 21]. Despite many types of underwater fishes, authors have collected image data of five marine fishes summarized in Table 2 from the Fish Recognition Ground Truth dataset. In the presented technique, 2000 images of each type of fish are retrieved contributing to a larger dataset of 10,000 images.

3.2 Fish image pre-processing

Underwater studies and exploration have gained pace in the recent years and the challenging illumination needs to be addressed using computer vision-based algorithmic designs. These images generally suffer due to light scattering and noise leading to low contrast and blur images. In 2017, Xie et al. [22] had implemented a similar

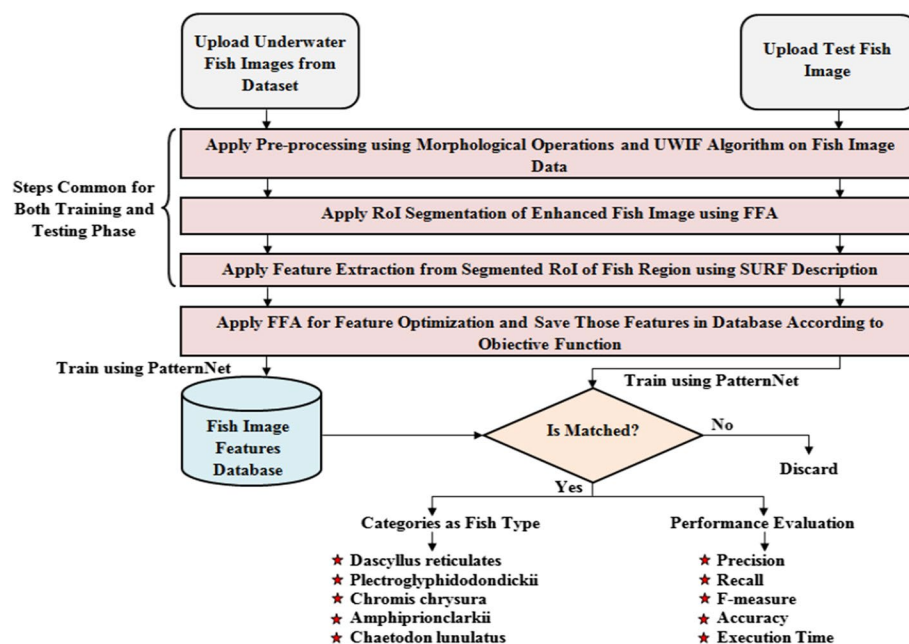


Fig. 2 Flowchart of Methodology

Table 1 Survey of various fish species classification models

Author	Methodology	Benefits	Drawbacks
Fouad et al. [10]	Scale Invariant Feature Transform (SIFT) along with Speeded up Robust Features (SURF) followed by Support Vector machine (SVM)	Experimentation had illustrated that 94.4% detection accuracy was attained by the fish classification utilizing SURF and SVM	The presented work was not utilized in a real-time environment
Qin et al. [17]	Framework centered upon Principal Component Analysis (PCA)	SVM was utilized as a machine learning classifier that illustrated 98.64% classification accuracy against a real-world fish recognition dataset	The filters of the proffered system were not as nice
Siddiqui et al. [18]	CNN	For the fish captured by the Western Australia coast, this strategy depicted 94.3% classification accuracy	In this work, the training data used was limited
Lakshmi et al. [20]	Gaussian mixture model + SURF	Utilizing multiclass-SVM (MSVM), the testing was executed which exhibited 88.9% classification accuracy	High false positives were possessed by this method
Jalal et al. [21]	YOLO deep neural network	Fish detection F-scores of 95.47% along with 91.2% were attained by this method, whilst fish species classification accuracies of 91.64% along with 79.8% were obtained on datasets correspondingly	More computational power was utilized by the proposed system since it comprised a complex machine learning tool when contrasted to traditional computer vision along with image processing approaches
Villon et al. [19]	Few Shot Learning (FSL)	This work found that the classic DL approach was outperformed by FSL in situations along with good classification accuracy was offered	The annotated images in this work were restricted

Table 2 The scientific names as well as the common names of the fish species

Sr. No	Scientific names	Common name
1	<i>Dascyllus reticulatus</i>	Two-stripe damselfish
2	<i>Plectroglyphidodon dickii</i>	Blackbar devil
3	<i>Chromis chrysura</i>	Stout chromis
4	<i>Amphiprion clarkii</i>	Clark's anemonefish
5	<i>Chaetodon lunulatus</i>	Oval butterflyfish

image processing for improving robotic navigation. In the present work, authors have addressed this issue by first applying mathematical operations followed by FFA-based optimization following Under-Water Image Enhancement that reflects instrumental image processing results.

The morphological operations in the fish image pre-processing are as follows:

- *Color conversion* It converts the uploaded fish image to a gray scale image using following conversion relationship. Red has to offer 30%, Green has to contribute 59% which is larger in every '3' color, and Blue has to give 11% for reducing the

contribution of red color, augmenting the contribution of the green color, and putting blue color contribution in betwixt these two.

$$\text{Fish}_{\text{Img}}(\text{Gray}) = 0.299 * \text{Fish}_{\text{img}}(\text{R}) + 0.587 * \text{Fish}_{\text{img}}(\text{G}) + 0.114 * \text{Fish}_{\text{img}}(\text{B}) \quad (1)$$

where, $\text{Fish}_{\text{Img}}(\text{Gray})$ represents the gray scale image, $\text{Fish}_{\text{img}}(\text{R})$ represents red component of the image, $\text{Fish}_{\text{img}}(\text{G})$ signifies the image's green component and $\text{Fish}_{\text{img}}(\text{B})$ denotes the image's blue component.

- *Binarization* It is also called masking of the image. Here, the gray scale fish image is represented as a matrix of 0's and 1's forming a binary image. This step helps in precise feature extraction process in the following steps of the methodology. It is performed using following mathematical relation:

$$\text{Image}_{\text{binary}}(i,j) = 1, \text{ if } \text{Fish}_{\text{Img}}(\text{Gray}) \geq \text{Th}_{\text{value}} \quad (2)$$

$$\text{Image}_{\text{binary}}(i,j) = 0, \text{ if } \text{Fish}_{\text{Img}}(\text{Gray}) < \text{Th}_{\text{value}} \quad (3)$$

where, $\text{Image}_{\text{binary}}$ represents a binary image with 'i' and 'j' representing rows and columns of the image.

- *Thinning* Usually after binarization any of the morphological operation like erosion, dilation, opening or closing or their combination is performed to achieve the boundary. Like other morphological operators, a structuring element determines the behavior of the thinning operation. The extended type explained under the hit-and-miss transform are the binary structuring elements utilized for thinning. But in the proposed method authors used thinning operation for skeletonization of foreground. This operation helps in distinguishing overlapping foreground pixels in the fish image. It is performed by applying hit and miss transformation on the binary fish image converted in the last operation. Mathematically thinning operation is expressed as follows:

$$\text{Fish}_{\text{Img}}_{\text{thin}} = \text{Image}_{\text{binary}} - \text{Hit\&Miss}(\text{Image}_{\text{binary}}, J) \quad (4)$$

Where, $\text{Fish}_{\text{Img}}_{\text{thin}}$ is the fish image obtained after thinning operation, Hit&Miss is applied to calculate linear pixels present in binary image, $\text{Image}_{\text{binary}}$ and 'J' represents the structuring element. The subtraction is a logical subtraction defined by $X - Y = X \cap \text{NOT } Y$. These operations proved to be very advantageous to improve the underwater image quality along with the Under-Water Image Enhancement algorithm.

The above algorithm works by first identifying the size of the input fish image followed by compression to minimize the memory requirements of the system. Image is split into individual planes and plane wise image enhancement is performed. Finally, all the enhanced planes are concatenated to return enhanced fish image $EFish_{\text{img}}$.

3.3 Image segmentation

FFA is a meta-heuristic algorithm based on swarm intelligence of flies that helps in extracting the fish images from the whole image. Yang had introduced this algorithm that is based on the interaction among the fireflies using their flashing lights. It is believed that being unisexual each firefly can be attracted by other firefly and this attractiveness is proportional to the light illuminated by the firefly. The Cartesian distance plays a role in the attractiveness between two fireflies and it is proportional to the brightness which reduces with augmenting distance between fireflies. In a region, if lesser luminance is possessed by all the fireflies, then those fireflies will move arbitrarily in the dimensional search space till a firefly with brighter luminance is found. For guiding the search process, the brightness of a firefly is associated with the analytical form of the objective function allocated. The cost function plays a major role in the FA's overall performance (exploration time, speed of convergence, along with optimization accuracy), and it monitors the optimization search. The luminance of a firefly is deemed to be proportional to the value of the cost function (that is., $\text{luminance} = \text{objective function}$) for a maximization issue. This brightness in turn is governed by the objective function and its strength is inversely proportional to the distance between them. Hence, this nature inspired algorithm allows filter out the irrelevant data and extract the most relevant solution. In the present methodology, FFA is implemented for the fish image segmentation from the underwater images. Segmentation is actually an approach in which the required Region-of-Interest (RoI) is extracted from the whole region under consideration. In the present case, the RoI is the fish image that needs to be extracted from the whole underwater image that contains fish in addition to its surroundings. The enhanced fish image is followed by fish RoI segmentation using FireFly algorithm (FFA) fitness function. The step proved to be very advantageous to extract the exact fish mask from the underwater fish images.

The segmentation process starts with the determination of dimensions of the underwater fish image under considerations. This information extracted in the form of knowledge of rows and columns is used in subsequent steps. Based on the threshold value mask of the fish is prepared and fish RoI is extracted and segmented fish image is returned as $SFish_{img}$. The FFA had greatly improved the quality of segmentation. The difference in RoI selection and segmentation with and without implementation of FFA for underwater fish images is compared in Table 2 of Result section.

3.4 Feature Extraction

In literature, Scale Invariant Feature Transform (SIFT) and Speeded up Robust Feature (SURF) have been widely used for feature-based extraction works. In fact, SIFT is also a good candidate for feature selection like SURF algorithm and both work equally good for illumination changes image. However, in the present methodology SURF is implemented for feature selection stage. This is due to fact that SURF extracts features based on the square filter and covers a larger area as compared to SIFT that works by using Gaussian filter. This considerably increases the overall accuracy and speed of SURF for feature extraction. The SURF features are detailed as follows.

SURF is a scale and rotation invariant feature descriptor used for image processing. It acts as both detector and descriptor. The detector locates the interest points in the image and descriptor describes the features of the interest points and constructs the feature vectors. Interest points on the image are detected using determinant of approximate Hessian matrix. It detects the blob-like features in an image. The blob detection is used to find the difference in image properties such as brightness, color, and corner. Let $X(c, d)$ denotes the point in the detected image; the hessian matrix $[h(X, u)]$ is formulated as,

$$[h(X, u)] = \begin{bmatrix} R_{XX}(X, u) & R_{XY}(X, u) \\ R_{XY}(X, u) & R_{YY}(X, u) \end{bmatrix}$$

Here, $X(c, d)$ mentions the convolution of Gaussian second-order derivatives $\frac{\partial^2}{\partial X^2} G(u)$ with the image x at point X where,

$$G(u) = \frac{1}{2\pi u^2} \text{Exp} \left[\frac{-(X^2 + Y^2)}{2u^2} \right]$$

The second-order derivative is used to find the edge magnitude at each pixel region. To avoid the filter of same size iteratively on the result of previously filtered image layer, integral image, and box filter are used. These filters represent the smallest scale for the calculation of blob response maps and are represented as $B_{XX}(X, u)$, $B_{XY}(X, u)$, $B_{YY}(X, u)$. It is therefore specified by,

$$|h_{app}| = B_{xx}B_{yy} - (WD_{xy}^2)$$

where, W represents the weight for the energy conservation between the Gaussian kernels and the approximated Gaussian kernels, then interest point matching is done based on the distance between template images and the searching images.

The present section is dedicated for the identification of feature points of the segmented fish image.

The segmented underwater fish image is used as input in the above described algorithm. In an iterative manner, extreme points of the segmented fish image are determined that are considered for the identification of other key extreme points present in its vicinity. In case, any other extreme point is detected, the localization is resigned for that pixel area. The algorithm returns $FishF_{points}$ as the best feature points for input segmented fish image. In order to precisely, select the only features that belong to fish part FFA objective function is used to determine a threshold value which is used to refine the SURF results. Here, feature points are optimized using FFA algorithm to return optimized fish feature data. Firefly plays an important role in filtering the features that actually belongs to fish type to improve the overall performance of fish classification. Suppose feature set of fish $F_a = \{f_1, f_2, f_3, \dots, f_n\}$, where n is the number of features for F_a

The threshold value f_t for features that belong to fish F_a is represented by the mean of all the feature values that belongs to F_a .

$$f_t = \frac{f_1 + f_2 + f_3 + \dots + f_n}{n} // \text{means of features in } F_a$$

$$f_t = \left(\sum_{i=1}^n f_i \right) / n \quad (4)$$

The objective function implemented for the feature selection is defined as follows:

```

for each  $f_i$  in  $F_a$  // reads each feature point of the feature data set  $F_a$ 
  iff  $f_i > f_t \rightarrow$  Accept the feature as fish feature.
  iff  $f_i \leq f_t \left\{ \begin{array}{l} \text{Discard } f_i \\ f_i = f_t \end{array} \right. // \text{Discard or replace the feature}
Endfor$ 
```

The above equations represent the two possibilities that arise when the feature f_i does not belong to fish F_a . Now, either the feature f_i can be discarded or the feature f_i can be replaced by f_t . However, in the present methodology feature f_i is replaced by the threshold value f_t determined using FFA-based objective function. This is because of the fact that if the feature f_i is chosen to be discarded, it will decrease the feature dataset size. Therefore, to improve the overall efficiency of the approach the feature f_i is replaced with f_t .

3.5 Classification

The testing and classification process requires trained features database of underwater fish images. In the present section, PatternNet is applied on the optimized features returned by the last step. A '3' layered architecture that comprises the input layer, hidden layer, along with output layer is the PatternNet. Artificial input neurons are encompassed by the neural network's input layer, and it brings the initial data into the system for further processing via subsequent layers of artificial neurons. The very start of the workflow for the neural network is the input layer. A layer between input layers and output layers is the hidden layer, in which a set of weighted inputs is taken by artificial neurons and it generates output via an activation function. The last layer of neurons that generates given outputs for the program is the output layer. The optimized feature data is taken by the input layer as the training data. While assessing errors, feature data is propagated by the hidden layer functions in forward along with cross-validation by backpropagation. Two-fold classification outcomes are returned by the output layer. The ability of the convolutional layers in a CNN is leveraged by PatternNet, where every filter usually has a constant response to certain high-level visual patterns. Utilizing a specially designed fully connected layer, this property is utilized for discovering discriminative along with representative visual patterns. For finding a sparse combination of filters, a loss function is used, which possess strong responses to the patterns in images as of the target category and weak responses to images as of the remaining categories. One is in the form of type of fish recognized by the proposed design and the other is in the form of performance parameters determining the quality of classification. The optimized features are trained using PatternNet architecture and stored in the database.

The above PatternNet algorithm works with optimized training fish data, number of neurons and the target data. The epoch and neuron number along with employed technique are initialized and on the basis of trained features of fish images, the input test images are categorized into five fish categories. In the process, weights are adjusted

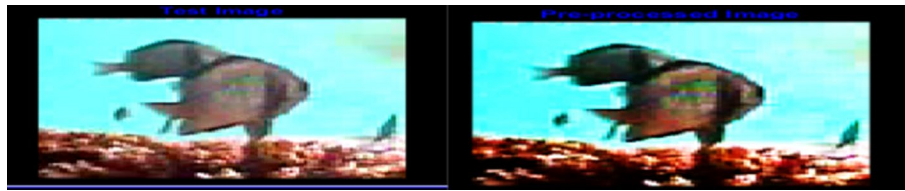


Fig. 3 Before and after Pre-processing of image

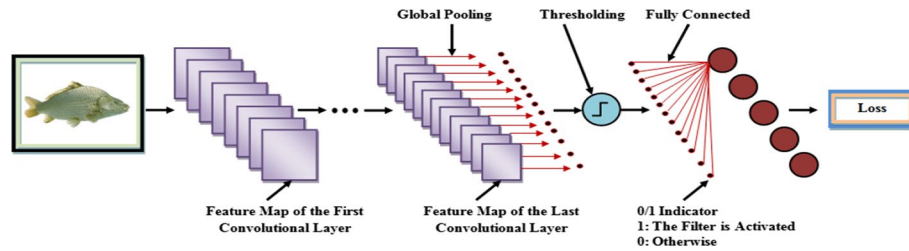


Fig. 4 General architecture of PatternNet, which contains input layers, several hidden layers, and output layers

to reach a desired output. Here, PatternNet is employed for both training and testing stages. The architecture of PatternNet is given in following figure (Figs. 3 and 4),

4 Results and Discussion

The working platform of the proposed work is MATLAB. The image quality of the underwater images is improved at each step to reflect an accurate fish classification. The outcomes of each processing step are summarized in Table 3. Row 4 and Row 5 displays the results of RoI selection and segmentation performed using threshold and morphological operations. However, Row 6 and Row 7 show the improvement in RoI selection and segmentation of fish images with the involvement of FFA-based optimization. A significant improvement in the fish image segmentation is observed by comparing before and after RoI selection and segmentation images. Next, SURF features for the respective selected areas of fish image are highlighted in Row 8.

4.1 Performance evaluation










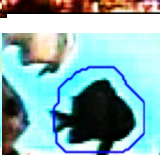

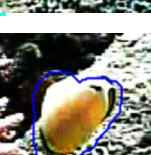
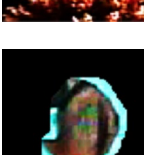
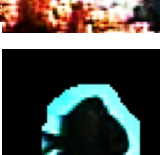




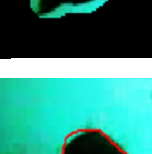

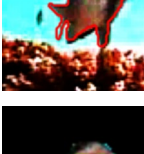

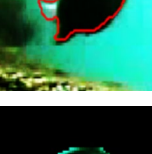

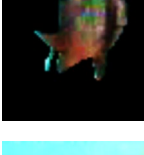
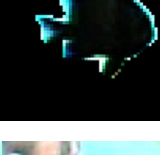

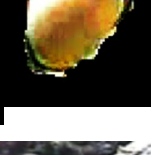
The correct recognition of the fishes is reflected in the form of confusion matrix. However, the efficiency of the proposed underwater fish classification technique is evaluated in terms of quality parameters concerning precision, recall, f-measure, error and accuracy. These are further calculated using following formulas:

- *Precision* The closeness of two or more measurements to one another is called precision.

$$\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}} \quad (5)$$

- *Recall* The total true positives that were found are called recall.

Table 3 The image analysis of several stages which are involved in the fish species classification

Fish Images		FI-1	FI-2	FI-3	FI-4
Original Underwater Fish Images					
Pre-processed Images using UWIE					
RoI Segmentation using Threshold and morphological operations	Selected RoI				
	Segmented RoI				
Improved ROI Segmentation with FFA-based optimization	Selected RoI				
	Segmented RoI				
Fish images with showing SURF Features					

$$\text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}} \quad (6)$$

- *F-measure* By taking the harmonic mean of the recall along with the precision, F-measure is assessed.

$$F - \text{measure} = 2 * \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (7)$$

- *Accuracy* The measure of correctness of the value in correlation with the information is called accuracy.

$$\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TruePositive} + \text{TrueNegative} + \text{FalsePositive} + \text{FalseNegative}} \quad (8)$$

The parametric values obtained for precision, recall, f-measure, accuracy and the required execution time using UWIE algorithm and FFA optimized UWIE algorithm are compared in Table 4. For this comparison various numbers of images are considered that are varied from 10 to 10,000. The proposed UWIE algorithm without optimization technique attains the precision of 0.86, recall of 0.8, F-measure of 0.82, accuracy of 98.09%, and execution time of 2.719 s. Likewise, the UWIE algorithm with FFA optimization technique achieves precision rate of 0.87, recall of 0.83, F-measure of 0.85, accuracy of 99%, and execution time of 3.64 s. Hence, it is observed that the optimization has improved the quality of classification over a range of fish images used in the present technique.

Precision evaluation of the proposed work using UWIE technique with and without FFA-based optimization is compared in Fig. 5. The number of fish image varying from 10 to 10,000 is plotted on X-axis against the precision values on Y-axis observed using UWIE alone and UWIE with FFA. For a smaller sample size of 50 fish images the precision of both cases is found to be below 80%; however it increases to 0.95 and 0.97 as the number of samples is increased to 10,000 using UWIE and UWIE with FFA, respectively. Overall, 1.5% increase in the average precision has been observed with the involvement of FFA-based optimization at image processing stage.

In the present scenario, recall is used to reflect the sensitivity of the employed technique to predict most relevant outcomes. Figure 6 shows that recall increases with an augment in the number of image samples. This means that as the sample size increases the sensitivity of the employed techniques increases; however, it is found higher using UWIE with FFA as compared to UWIE alone. Overall, it is observed that FFA-based optimization resulted in 3.45% improved recall value.

The combined effect of precision and recall of the proposed techniques is observed using f-measure values. Figure 7 shows that f-measure increase from 0.714 to 0.94 using UWIE while using FFA with UWIE increases f-measure from 0.734 to 0.975. This means that optimization resulted in 2.5% increased f-measure as a combined effect when 10,000 fish image samples were used.

Figure 8 compares the accuracy of fish classification for a sample size ranging from 10 to 10,000 fish images. Accuracy of 97.242% is observed for smaller sample size of 10 fish images that increases to 98.188% with the involvement of FFA-based optimization. The involvement of velocity parameter increases the classification accuracy to a higher level as compared to using UWIE alone. The accuracy against 10,000 image sample is observed to be 98.845% and 99.008% using UWIE and UWIE with FFA, respectively. However, an average increase of 0.916% has been achieved with the involvement of optimization technique. Involvement of optimization strategy not

Table 4 Performance Evaluation of the proposed methodology in terms of precision, recall, F-measure, accuracy and execution time using UWIE with FFA and without FFA

Number of Images	Precision		Recall		F-measure		Accuracy (%)		Execution Time (s)	
	UWIE	UWIE + FFA	UWIE	UWIE + FFA	UWIE	UWIE + FFA	UWIE	UWIE + FFA	UWIE	UWIE + FFA
10	0.74	0.76	0.69	0.71	0.714	0.734	97.242	98.188	1.954	2.234
20	0.76	0.78	0.71	0.73	0.734	0.754	97.348	98.394	2.018	2.415
50	0.78	0.79	0.72	0.74	0.749	0.764	97.486	98.495	2.114	2.546
100	0.82	0.83	0.73	0.75	0.772	0.788	97.846	98.592	2.175	3.014
200	0.84	0.85	0.75	0.79	0.792	0.819	97.974	98.994	2.354	3.345
500	0.87	0.89	0.78	0.8	0.823	0.843	98.214	99.187	2.548	3.784
1000	0.91	0.91	0.82	0.85	0.863	0.879	98.248	99.257	2.842	3.958
1500	0.92	0.94	0.87	0.91	0.894	0.925	98.485	99.348	3.013	4.162
2000	0.93	0.95	0.89	0.95	0.910	0.950	98.584	99.445	3.242	4.586
5000	0.94	0.96	0.91	0.97	0.925	0.965	98.745	99.574	3.674	4.887
10,000	0.95	0.97	0.93	0.98	0.940	0.975	98.845	99.618	3.975	5.174

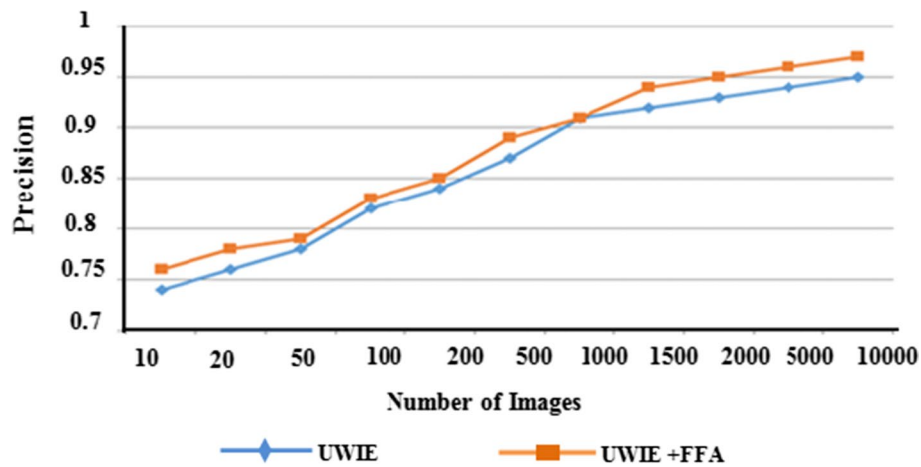


Fig. 5 Graphical representation of precision evaluation

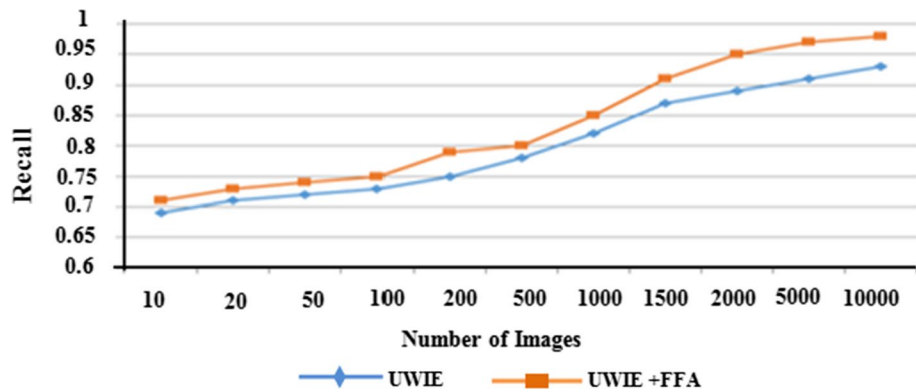


Fig. 6 Graphical representation of recall evaluation

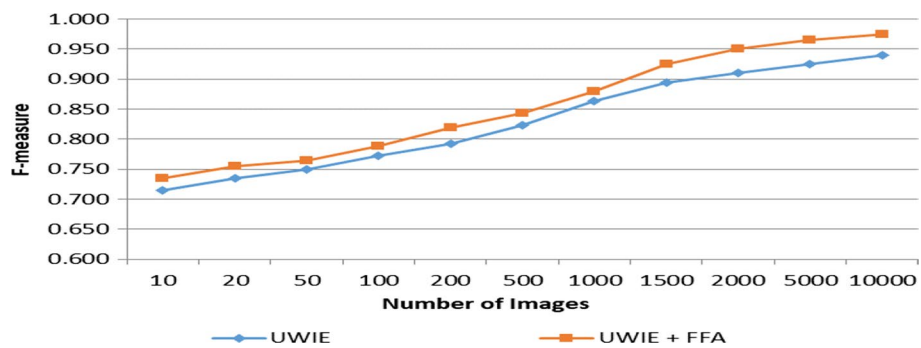


Fig. 7 Graphical representation of F-measure evaluation

only increased the quality parameters of classification but also increased the execution time of the proposed technique.

Figure 9 compares the execution time used for performing classification over the sample size ranging from 10 to 10,000 fish image samples. It is observed that for small

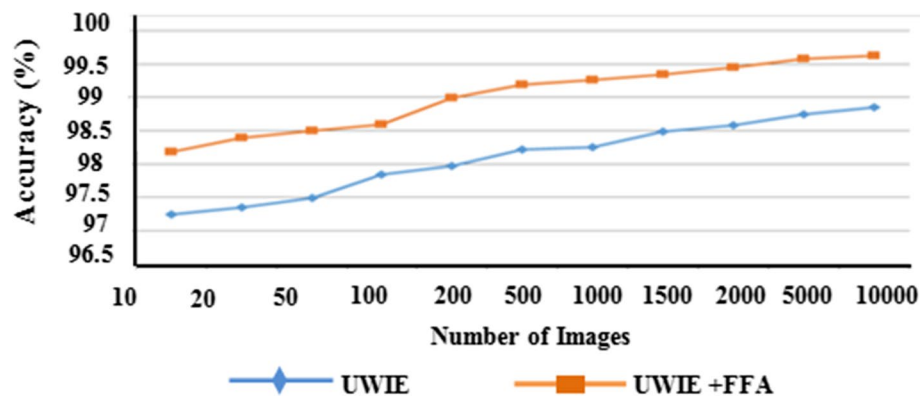


Fig. 8 Graphical representation of accuracy evaluation

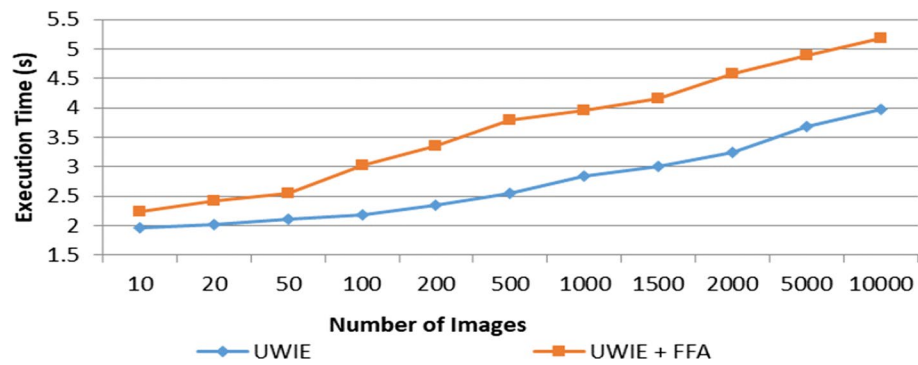


Fig. 9 Graphical representation of execution time evaluation

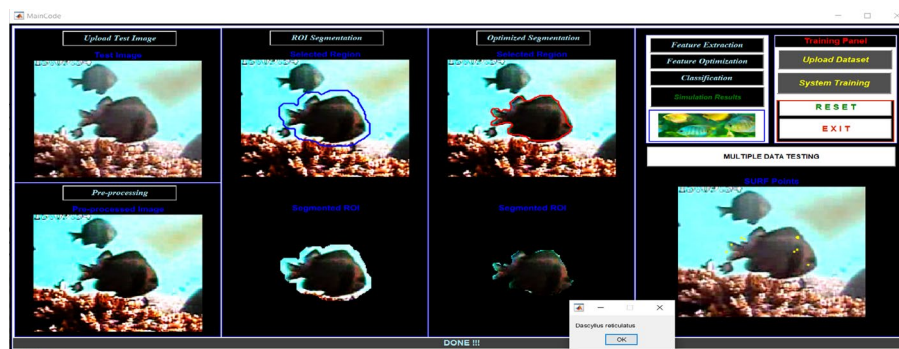


Fig. 10 Testing results execution

sample size lesser execution time was required that increases considerably as the sample size gets larger. However, using UWIE alone time required is from 1.954 to 3.975 s that increases from 2.234 to 5.174 s with the involvement of FFA optimizations.

4.2 Performance evaluation with classification time and accuracy rate

The performance of the proposed-Framework has also been evaluated by analyzing execution time taken for classification of the tested images to various five categories.

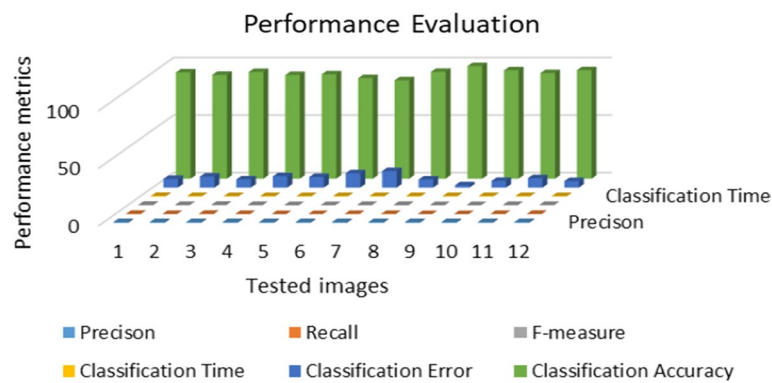


Fig. 11 Performance Evaluation of proposed-framework with respect to Precision, Recall, F-measure, Classification Time, Accuracy rate in Classification and classification error

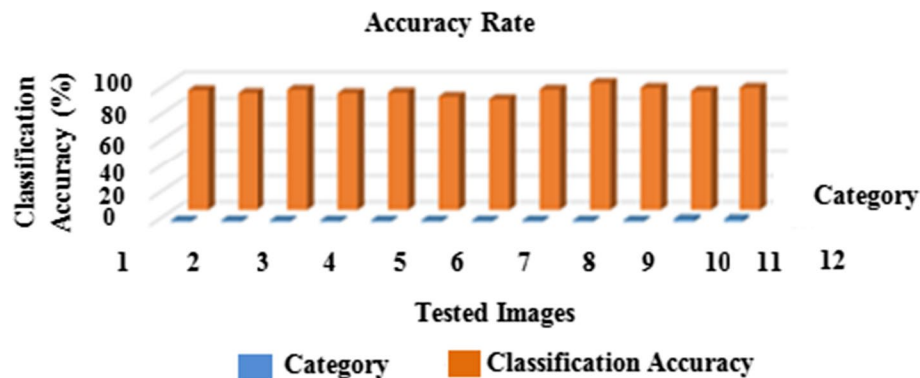


Fig. 12 Graphical representation of accuracy rate evaluation

The feature optimization for SURF-based feature extraction to improve PatternNet-based classification of underwater fish images into five distinct fish classes is implemented in the method.

Figure 10 illustrates the execution of the input tested images, to obtain the features selection of the Fish-images through the sequential process. The test images are uploaded in the panel, followed by pre-processing methods. The regions of the images undergo segmentation and ROI segmented image samples are depicted in the figure above. The optimized ROI Segmented images, is applied with SURF-Speed-up Robust-feature algorithm, to obtain the features extracted from the segmented ROI sections. The Features extracted are represented as SURF-points.

Figure 11 illustrates the performance metrics evaluation of all the 12 tested images with respect to Precision factor, Recall value, F-measure value, Classification Time, Accuracy rate in Classification and classification error. The twelve sets of tested Fish images are subjected to proposed-framework, and the outcomes are depicted in the figure. From the representation above, it is found that the classification of categorized outcomes exhibits higher accuracy rate in classification of fish-species to categories 1 and category 2. The precision values of the proposed-framework also found to be higher in all the images classification.

Table 5 Comparison of classification accuracy for various categories of fish

Category	Classification accuracy
1	92.42512892
1	90.37327301
1	92.88365387
1	90.19030273
1	90.77441325
1	87.50047688
1	85.59604354
1	92.9655511
1	97.96288342
1	94.20217807
2	91.85289205
2	94.33057445

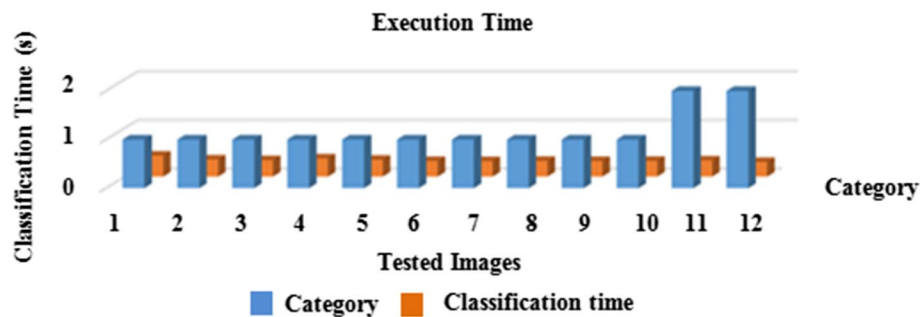
**Fig. 13** Graphical representation of classification time evaluation

Figure 12 represents the variations of classification accuracy rates of twelve tested images. The input tested images are subjected for performance evaluation, in assessing the accuracy percentage of the proposed-framework for tested set of twelve images. It is determined that the classification accuracy of various tested-images is higher throughout in all images. The twelve tested images are analyzed with the patterns recognized in the trained 38 image samples, resulting in the classification of categorized images.

Table 5 represents various rates of accuracy metric, in classifying the given tested images obtained from the segmentation process. The various levels of accuracy rates for all obtained images are determined in this table. The accuracy rates of categorized images range up to 98 percentage, specifying higher range in accuracy percentage. This high rate of accuracy parameter exhibits higher efficiency of proposed-framework.

Figure 13 illustrates graphical representation, in evaluating the execution time of proposed-framework classification. The execution time with respect to the tested-images has been plotted in the graph. The classification execution time is determined for various tested images and made in comparison to the various images classification process. From the figure, it is depicted that the execution time in

Table 6 Representation of classification time required for classifying different categories of fish

Category	Classification time
1	0.4297826
1	0.3513445
1	0.3377436
1	0.3735447
1	0.3467464
1	0.3203421
1	0.3170541
1	0.3225374
1	0.3203741
1	0.3244439
2	0.3316675
2	0.300798

Table 7 Accuracy Comparison of the proposed framework and the existing works

Reference work	Techniques used	Accuracy
Proposed Work	UWIE + SURF + PatternNet	98.87
Fouad et al. [10]	SIFT + SURF and SVM	94.4
Siddiqui et al. [18]	CNN followed by SVM	94.3
Vilon et al. [19]	CNN with decision rule	94.9

classifying twelve images seems to be considerably lesser, showing the efficiency of the proposed-framework.

The execution time taken for entire Segmentation, features-extraction and optimization proceeded by the classification process is evaluated for tested-fish images in Table 6. The Classification time taken for categorizing the fish-type ranged from 0.3 to 0.4 values. The execution time hence seems to be lesser for all the categorized tested images, which implies the performance efficiency of the proposed-framework.

4.3 Comparison against existing studies

The performance of the proposed work is also evaluated against the demonstrated fish determination accuracy of various existing works summarized in Table 7. SIFT and SURF algorithms were employed for feature extraction of fish images by Fouad and co-researchers to correctly identify Nile Tilapia fishes using SVM classifier. This combination resulted in accuracy of 94.3% [10]. In 2018, Siddiqui employed a CNN-based feature extraction followed by SVM-based classification to reflect an accuracy of 94.3% [25]. However, another researcher Vilon and researchers employed decision rules along with CNN to determine the fish type from fish image dataset to achieve accuracy of 94.9% [18].

It is observed that various combinations have been employed by numerous researchers in order to correctly identify fishes and categorize them into various species. Figure 14 graphically compares the demonstrated accuracies of proposed work against the existing works. It is observed that UWIE algorithm followed by FFA for fish image segmentation had enormously improved the segmentation ability of the proposed work. Further, FFA-based

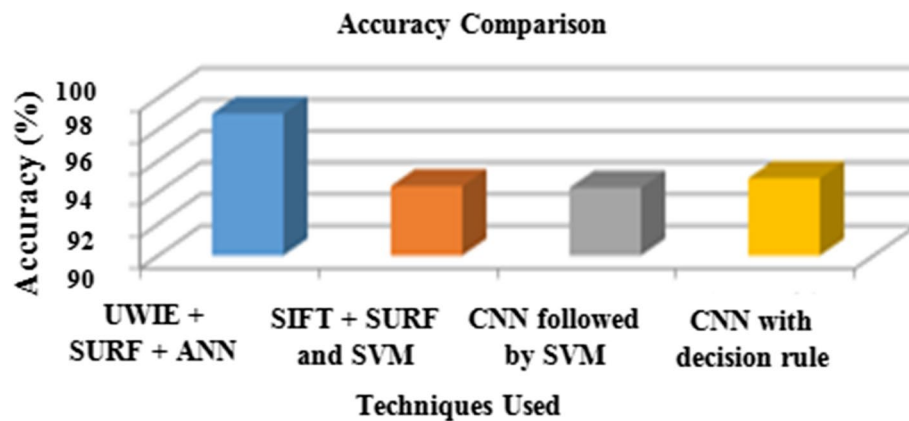


Fig. 14 Comparative analysis of accuracy rate of several techniques

optimization of extracted fish image features added to increased classification accuracy using PatternNet classifier. Overall, proposed work outperformed existing work of Fouad et al., Siddiqui et al., and Vilon et al. by 4.6%, 4.5% and 4.1%, respectively.

5 Conclusion

In this paper, an efficient classification algorithm PatternNet technique is implemented in classification of Fish images to various categories. The morphological operations such as poor-intensity modifications, noise removal sharpening of images, edge-detection operations, de-blurring process facilitates to handle the images in Feature selection process. In this study, UWIE-techniques have been employed, enhancing the FFA-Firefly algorithm velocity functions in improvising the fish-classification. Pre-processing techniques remove the intensity variations in the images followed by Feature-extraction process. The features are selected by SURF-Speedup Robust-Feature algorithm. The feature optimization for SURF-based feature extraction to improve PatternNet-based classification of underwater fish images into five distinct fish classes is applied for the categorization of Fish-species, PatternNet technique is employed in classifying 10,000 marine fish-images to five distinct categories. The performance evaluation of the proposed-Framework is assessed concerning precision value, F-measure value, Recall factors, Classification time with respect to implementation, and accuracy rate evaluation. The classification accuracy and all the performance metrics values seem to be increasing while UWE integrated with Fire-Fly algorithm. The accuracy rate of this proposed-framework seems to attain 98% with lesser average computation time of 3.64 s. Furthermore, the proposed model revealed competitive results even for the blurred and illuminated images which reflect the robustness of the proposed model. Hence, with these inferences of the results, this depicted the better efficiency of the proposed-framework, and it overtakes the existing Fish classification and Recognition techniques.

5.1 Future scope

In future work, the proposed approach will be evaluated against a larger database, especially video datasets acquired in the unconstrained underwater environment. Moreover,

the proposed model will be extended to estimate fish size, weight, and age which are important for stock assessment and management.

Author contributions

Both authors read and approved the final manuscript.

Declarations

Competing interests

Both authors declare that they have no competing interest.

Received: 29 July 2022 Accepted: 18 November 2022

Published online: 02 December 2022

References

1. E.A. Awalludin, T.N.T. Arsad, W.H.W. Yussof, A review on image processing techniques for fisheries application. in *Journal of Physics: Conference Series*, vol. 1529, no. 5, (IOP Publishing, 2020), p.052031
2. M.K. Dutta, A. Issac, N. Minhas, B. Sarkar, Image processing based method to assess fish quality and freshness. *J. Food Eng.* **177**, 50–58 (2016)
3. B.J. Boom, J. He, S. Palazzo, P.X. Huang, C. Beyan, H.M. Chou, F.P. Lin, C. Spampinato, R.B. Fisher, A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage. *Ecol. Inform.* **23**, 83–97 (2014)
4. I. Aliyu, J.K. Gana, A.M. Aibinu, J. Agajo, A.M. Orire, T.A. Folorunso, M.A. Adegboye, A proposed fish counting algorithm using digital image processing technique. *ATBU J. Sci. Technol. Educ.* **5**(1), 1–11 (2017)
5. V. Allken, N.O. Handegard, S. Rosen, T. Schreyeck, T. Mahiout, K. Malde, Fish species identification using a convolutional neural network trained on synthetic data. *ICES J. Mar. Sci.* **76**(1), 342–349 (2019)
6. S. Marini, E. Fanelli, V. Sbragaglia, E. Azzurro, J. Del Rio Fernandez, J. Aguzzi, Tracking fish abundance by underwater image recognition. *Sci. Rep.* **8**(1), 1–12 (2018)
7. J. Le, L. Xu, An automated fish counting algorithm in aquaculture based on image processing. in *2016 International Forum on Mechanical, Control and Automation (IFMCA 2016)* (Atlantis Press, 2017)
8. M.K. Alsmadi, K.B. Omar, S.A.M. Noah, Fish classification based on robust features extraction from color signature using back-propagation classifier. *J. Comput. Sci.* **7**(1), 52 (2011)
9. FAO, *The state of world fisheries and aquaculture 2020* (Sustainability in action, Rome, 2020). <https://doi.org/10.4060/ca9229en>
10. M.M.M. Fouad, H.M. Zawbaa, N. El-Bendary, A.E. Hassanien, Automatic Nile tilapia fish classification approach using machine learning techniques. in *13th International Conference on Hybrid Intelligent Systems (HIS 2013)*. IEEE (2013), pp.173–178
11. R. Sethi, I. Sreedevi, Adaptive enhancement of underwater images using multi-objective PSO. *Multimed. Tools Appl.* **78**(22), 31823–31845 (2019). <https://doi.org/10.1007/s11042-019-07938-x>
12. Osman NoorulShuhadah, Mustaffa Mas Rina, Doraisamy Shyamala C., MadzinHizmahwati, "Content-based Image Retrieval for Fish based on Extended Zernike Moments-Local Directional Pattern-Huecolor Space." *International Journal of Innovative Technology and Exploring Engineering*, no. 8, vol. 8 (2019):173:183.
13. M. Ravanbakhsh, M. Shortis, F. Shaifat, A.S. Mian, E. Harvey, J. Seager, An application of shape-based level sets to fish detection in underwater images. in *GSR* (2014)
14. J. Hu, D. Li, Q. Duan, Y. Han, G. Chen, X. Si, Fish species classification by color, texture and multi-class support vector machine using computer vision. *Comput. Electron. Agric.* **88**, 133–140 (2012)
15. M.-C. Chuang, J.-N. Hwang, K. Williams, A feature learning and object recognition framework for underwater fish images. *IEEE Trans. Image Process.* **25**(4), 1862–1872 (2016)
16. A. Salman, S.A. Siddiqui, F. Shafait, A. Mian, M.R. Shortis, K. Khurshid, A. Ulges, U. Schwanecke, Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system. *ICES J. Mar. Sci.* (2019). <https://doi.org/10.1093/icesjms/fsz025>
17. H. Qin, X. Li, J. Liang, Y. Peng, C. Zhang, DeepFish: accurate underwater live fish recognition with a deep architecture. *Neurocomputing* **187**, 49–58 (2016)
18. S. Villon, D. Mouillot, M. Chaumont, E.S. Darling, G. Subsol, T. Claverie, S. Villéger, A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Eco. Inform.* **48**, 238–244 (2018)
19. G.D. Lakshmi, K.R. Krishnan, Analyzing underwater videos for fish detection, counting and classification. in *International Conference On Computational Vision and Bio Inspired Computing*, (Springer, Cham, 2019) pp.431–441
20. A. Jalal, A. Salman, A. Mian, M. Shortis, F. Shafait, Fish detection and species classification in underwater environments using deep learning with temporal information. *Eco. Inform.* **57**, 1–13 (2020)
21. S. Villon, C. Iovan, M. Mangeas, T. Claverie, D. Mouillot, S. Villéger, L. Vigliola, Automatic underwater fish species classification with limited data using few-shot learning. *Eco. Inform.* **63**, 1–6 (2021)
22. Fish Recognition Ground Truth database, Available Online at: <http://groups.inf.ed.ac.uk/f4k/GROUNDTRUTH/RECOG/>.
23. B.J. Boom, P.X. Huang, J. He, R.B. Fisher, Supporting ground-truth annotation of image datasets using clustering. in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, IEEE (2012), pp.1542–1545

24. K. Xie, W. Pan, Xu. Suxia, An underwater image enhancement algorithm for environment recognition and robot navigation. *Robotics* **7**(1), 14 (2018)
25. S.A. Siddiqui, A. Salman, M.I. Malik, F. Shafait, A. Mian, M.R. Shortis, E.S. Harvey, Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES J. Mar. Sci.* **75**(1), 37–389 (2018)

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
