

# Fish Classification using Saliency Detection Depending on Shape and Texture

Rafsun Jany Arman<sup>1</sup>, Monowar Hossain<sup>1</sup>, Sabir Hossain<sup>2</sup>

<sup>1</sup> Bangabandhu Sheikh Mujibur Rahman Science and Technology University,  
Computer Science and Engineering,  
Bangladesh

<sup>2</sup> Chittagong University of Engineering and Technology,  
Computer Science and Engineering,  
Bangladesh

{armanrafsunjany, murad0904045}@gmail.com, sabir.cse@cuet.ac.bd

**Abstract.** Classification of fishes becomes important after the advancement of machine learning. As fishes play a vital role in the economy of Bangladesh, a proper monitoring system will maximize the cultivation. It will also contribute to the overall economy. Therefore, here introduce a system that can detect the fishes and compare various methods with explanations to understand the selected methods. This paper have considered 5 categories of local fishes of Bangladesh in the dataset. The technique consists of preprocessing with segmentation, feature descriptor, and ensembles to produce the final result. U<sup>2</sup>-net is used in the preprocessing layer to obtain two types of features namely shaped images and colored images with removed backgrounds. To get the features, we have used a histogram of oriented gradient (HOG) and an ensemble layer is used for classification purposes. Experimental results illustrate the accuracy of 99.77% for the first ensemble and 100% for the second ensemble layer on our dataset of 2678 fishes of 5 distinguishing classes. Various layers were used to compare the predicted results using different performance metrics.

**Keywords.** U<sup>2</sup>-net, hog, knn, SVM, logistic regression, decision tree, fish classification, segmentation, salient object detection.

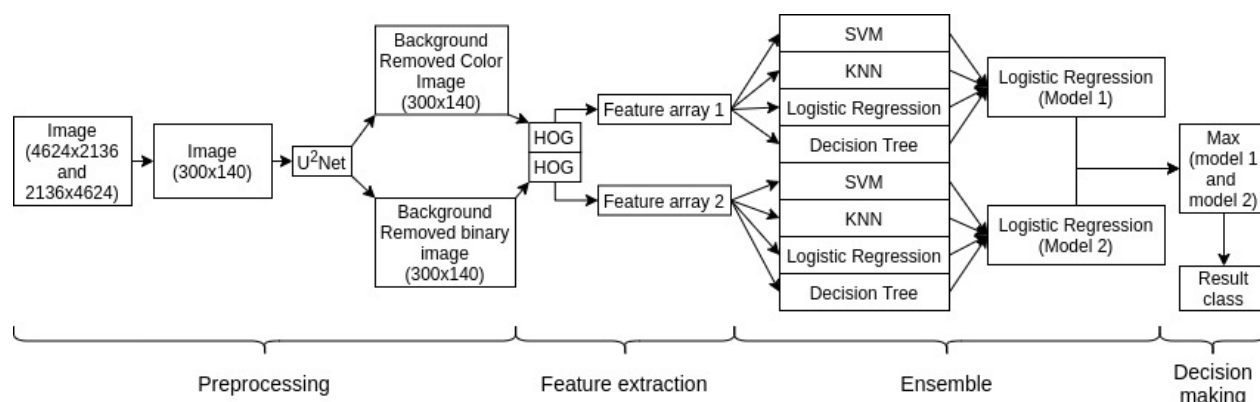
## 1 Introduction

As a delta state, Bangladesh has a huge number of rivers across the various parts of the country. Since the production of fish contributes to the

everyday life of millions of people in Bangladesh, this is considered to be the second most valuable agricultural crop in the whole country [13]. Amongst 32,000 species of fish worldwide, almost 40% of those species live in freshwater, as in our country, marine and inland fish (in fresh-waters and brackish waters) have a huge number of 401 and 251 species respectively [7]. So, we cannot ignore the importance of fish in our people's lives.

Hence, there is a need to have a fish processing unit that will help to produce a comparatively better classification system that helps the processing unit to collect the fish data on a conveyor belt in any kind of fish processing company. Along with that, there is a need to collect the shape info of the various classes of fish to make a processing system that processes or packs the same size fish. But there is a problem with the background. Various types of background make classification more difficult.

Therefore, spectral characteristics are used [23] to overcome the underwater turbulence and other noises because underwater there are environmental variations in luminosity, fish camouflage, dynamic backgrounds, water murkiness, low resolution, shape deformations of swimming fish, and subtle variations between some fish species [16]. We need a background independent fish classification (FC) using segmentation [24, 1]



**Fig. 1.** Four layers of proposed model, namely (i) preprocessing, (ii) feature extraction, (iii) ensemble, and (iv) decision making

depending on shape, texture, color. Moreover, some techniques involve dividing a fish into several parts like fish head, body, tail [6, 2]. Salp Swarm Algorithm (SSA) and threshold Otsu's method also produce a satisfactory FC result [15].

Boundaries of segmented regions and the contour extraction are improved by a proposed system using the median-cut algorithm [22]. However, many of the recent papers propose a feature descriptor using a transfer learning approach using pre-trained models like VGG16 [3, 9] and AlexNet [2].

After extracting the features various deep learning algorithms are used to classify like Artificial Neural Network (ANN) [24, 17, 12], Convolution Neural Network (CNN) [3, 18, 21, 8, 30, 27, 11], Deep Learning Network (DLN) [2]. Many of the recent papers suggested FC using machine learning techniques like Support Vector Machine (SVM) [28] and it is also used with the feature descriptor like Hybrid Linear Binary Pattern (HLBP) as classifier [29], Nearest Neighbors (KNN) [19], Decision Tree (DT) [28], Naive Bayesian as a fusion layer of DLN [2].

However, segmentation removes the maximum number of unnecessary features and collects only the regions of interest. So for the mentioned reasons, the suggested model use  $u^2$ -net to remove the background and a feature extractor like HOG to extract the features from the selected

region. Then a classification model is used to classify the classes of fish.

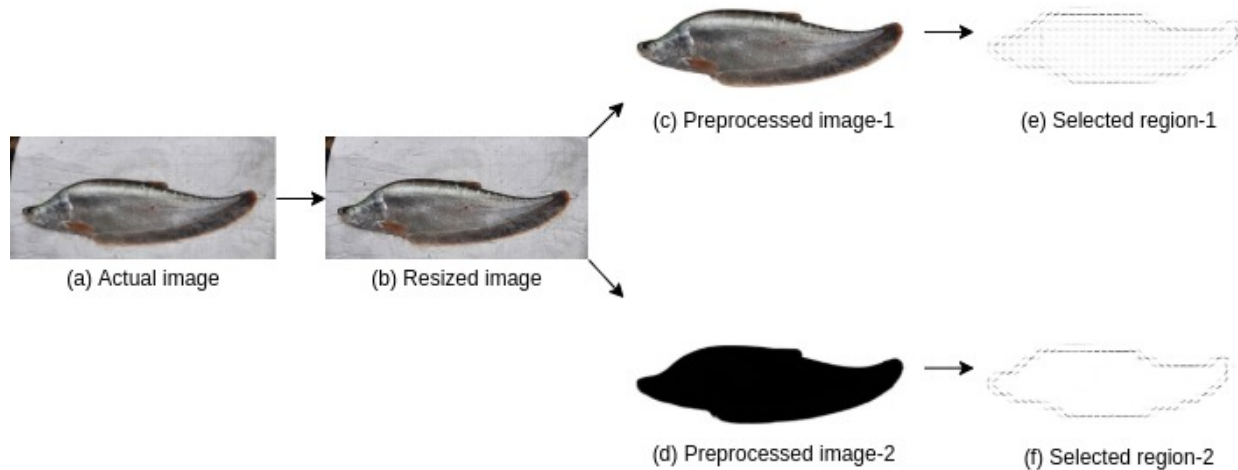
This paper significantly contributes to the following aspects, (i) Efficiently classify fishes by removing variant background using transfer learning techniques like  $u^2$ -Net, (ii) Extract features depending on shape and color images by HOG, (iii) Prepare a new fish dataset containing 2,678 samples of five classes.

The following parts of the article are distributed as, Section-2 describes the overall technique or methodology of the detailed discussion on the proposed classifier, Section-3 and 4 illustrate the dataset preparation techniques and the result analysis respectively, and Section-5 concludes the article itself.

## 2 Methodology

The proposed technique has four layered structures described in fig. 1. Preprocessing stage mainly preprocessed the images. It divides those images into two groups as follows: (i) focusing on shape (ii) colored images (background removed). These groups are named as background removed binary image and background removed color image respectively in fig. 1.

HOG is used to generate the feature arrays of two types of images. The details are described in further sections. So, there will be two types of feature arrays and two ensemble stacking



**Fig. 2.** Image preprocessing and feature extraction

classifiers are needed to classify the features individually (fig. 1). Finally, the decision-making layer uses the previous two ensembles to make decisions depending on the maximum prediction rate.

## 2.1 Preprocessing

Firstly, an actual image of the fish from our dataset has been taken as fig. 2(a). But the real image is huge in size (4624x2136) so that it requires resizing for computational purposes. Images are resized into 140x300 with an unchanged aspect ratio illustrated in fig. 2(b).

There is a problem with variant backgrounds on FC so that the classification task can't identify the true area of interest on an image because any image contains a huge area that does not play any kind of role on FC.

To overcome the issue, the proposed method adopted a deep learning algorithm, called U<sup>2</sup>-Net [26] as a transfer learning approach. It captures contextual information on a different scale from an image.

So after applying the method, it produces the mask image like fig. 2(d), and the mask is then used to remove the background from an actual image as fig. 2(c).

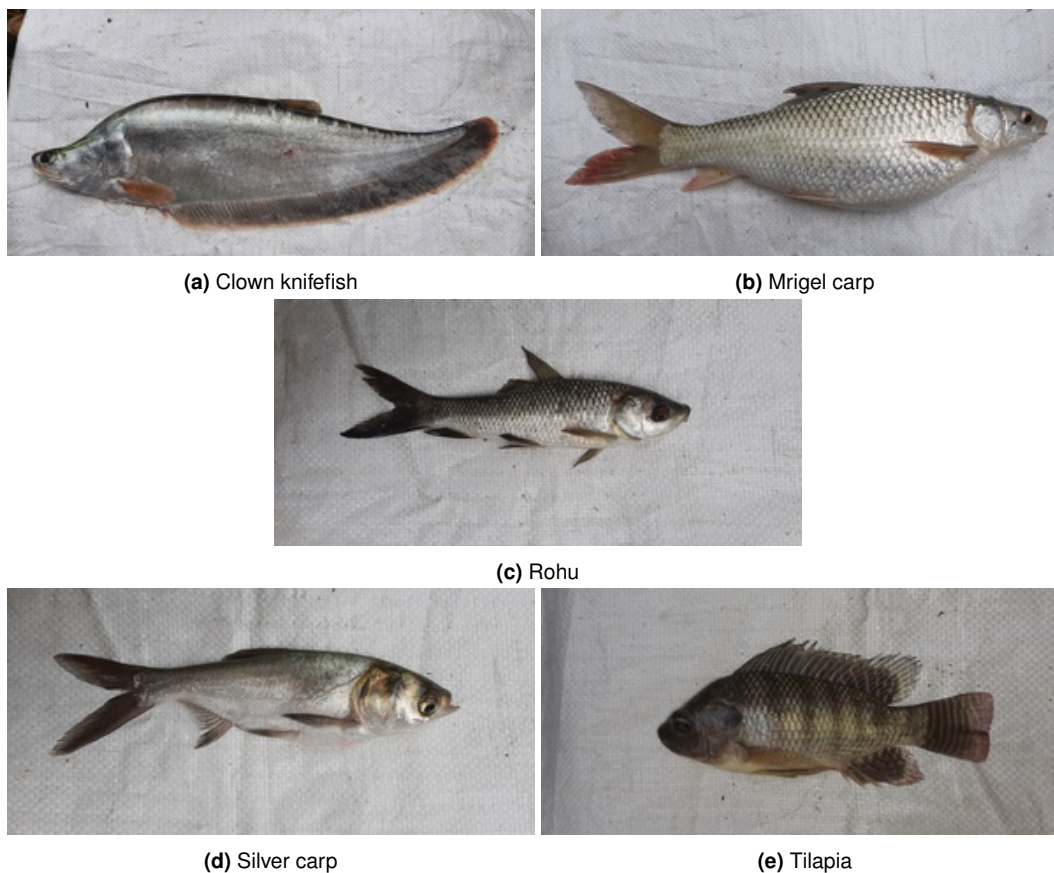
## 2.2 Feature Extraction

After getting fig. 2(c) and fig. 2(d), HOG [10] descriptor extracts the features. This algorithm basically focuses on the magnitude as well as on the direction. In addition to that, it breaks an image into several parts to capture the magnitude and the orientation. It produces not only the edge value but also the direction of the edge. So that, the proposed method suggested applying the HOG on both fig. 2(c) and fig. 2(d).

The resulting image from fig. 2(c) to fig. 2(e) demonstrates the inner edges along with the outer edges. Since the source is an RGB image, the output contains the edges of the fish body(inner edges) and shape(outer edges). But on the other hand, fig. 2(f) captures the outer edges from fig. 2(d). As fig. 2(d) is only a black and white image so that there is no inner edge left on the resulting image. However, the preprocessing phase produces two types of features illustrated above namely Feature array 1 and 2 (fig. 1) which are being used to classify the fishes separately.

## 2.3 Ensemble for Classification

In the preprocessing stage, the maximum number of unnecessary parts of an image is removed by the salient object detection technique. This



**Fig. 3.** Describing the original five classes of fishes from dataset

means now ours have the feature arrays of selected region-1 and 2 described in fig. 2(e) and fig. 2(f) respectively with minimalist areas. These areas satisfy our goals of feature selection because features only contain the fish details. So the proposed model suggested using shallow machine learning approaches like SVM, KNN, Logistic Regression, and Decision Tree to create the ensemble classifier.

Ensemble classifier is a technique that is a composition of many individual homogeneous or heterogeneous classifiers. There is an ensemble model called a stacking classifier. It is made of two parts: (i) base learner which is used as a training layer and (ii) meta learner which is used as a decision layer on the stacking classifier. The suggested model proposed a stacking ensemble classifier with SVM, KNN, Logistic Regression,

and Decision Tree for the base learner. They produce four separate prediction results on the dataset and finally a Logistic Regression model for meta learners that decides the final result from the previously generated prediction results.

As there are two types of selected regions, two stacking ensembles are needed to train and test on both sets of regions. As illustrated in fig. 1 there are models 1 and 2 generated by the mentioned technique from feature arrays 1 and 2 respectively.

## 2.4 Decision Making

Model 1 and model 2 are then used to produce the result of the final decision layer. They both give the five different percentages of each fish class. So, there are 10 different prediction values from

**Table 1.** Details of the dataset

Local Name	Eng. Name	# Images	Split Size	
			Training	Testing
Chitol	Clown knifefish	610	408	202
Mrigel	Mrigal carp	616	412	204
Rui	Rohu	642	430	212
Silver carp	Silver carp	596	399	197
Telapia	Tilapia	214	143	71
Total		2678	1792	886

**Table 2.** Performance evaluation of individual machine learning models

Input Data	Classifier Name	Classification Accuracy	
		Training	Testing
Feature array 1	SVM	99.94%	99.66%
	KNN	99.83%	99.77%
	Logistic Regression	100%	99.66%
	Decision Tree	100%	91.094%
Feature array 2	SVM	100%	99.66%
	KNN	99.83%	99.66%
	Logistic Regression	100%	99.66%
	Decision Tree	100%	93.80%

**Table 3.** System details

System	Type	Details
CPU	Model name	Intel(R) Core(TM) i7-6700HQ CPU @ 2.60GHz
	Architecture	x86_64
	RAM	16GB
	OS	Ubuntu 20.04.3 LTS
	VGA compatible controller	Intel Corporation HD Graphics 530 (rev 06)
GPU	Model name	GeForce GTX 960M (rev a2)

**Table 4.** Performance evaluation of the Model 1 and 2

Name	Classification Accuracy	
	Training	Testing
Model 1	99.89%	99.77%
Model 2	100%	100%

both models 1 and 2. This layer performs a max voting approach to decide the finally generated fish class from those 10 probability values. Maximum probability is decided to be the final output class or result class.

### 3 Dataset Preparation

Dataset consists of 2678 numbers of images of five different fishes of Rohu, Mrigal carp, Silver carp, Clown knife fish, and Tilapia described in

table 1. The resolution of the images is 4624x2136. Images were captured with Samsung S5KGW1 censored camera. Every class is captured in a different position and collected from Bangladeshi local ponds. In fig. 3, there is a sample of images collected for the training and testing.

After preparing the dataset, it has been divided into train and test split with a 77% and 33% ratio respectively. To pick the random images for the train and test set, `sklearn.model_selection.train_test_split` is used. But there is one problem: if the whole dataset is divided into train and test sets, there will be a chance that an imbalance number of fish classes will be added into the train and test. So, to overcome this problem every class of fish is divided separately into train and test set as described in table 1.

## 4 Result Analysis

In table 3, presented the system environment. This system is used to evaluate the models. The next few subsections describe the analysis process.

### 4.1 Base Learners Accuracy

Recommended models (models 1 and 2) have two layers of base learners. So, observation is conducted through the base learners to evaluate the inner structure of the stacking models as follows in table 2. Observed accuracy shows that the classifiers make an outstanding performance on the created dataset both on training and testing due to using the extra feature deduction layer as the paper suggested.

### 4.2 Final Model Accuracy

On table 2, individual classifiers show good results. Another meta learner is used to predict the previous base learner's result. It gives the consistent and the final accuracy of each feature array. Hence, a logistic regression classifier is used to supervise the result of the previous layers. So, there need to be two results of accuracy both for models 1 and 2. And they produce a better accuracy described in table 4. Besides, confusion matrix-based evaluation is conducted on models 1

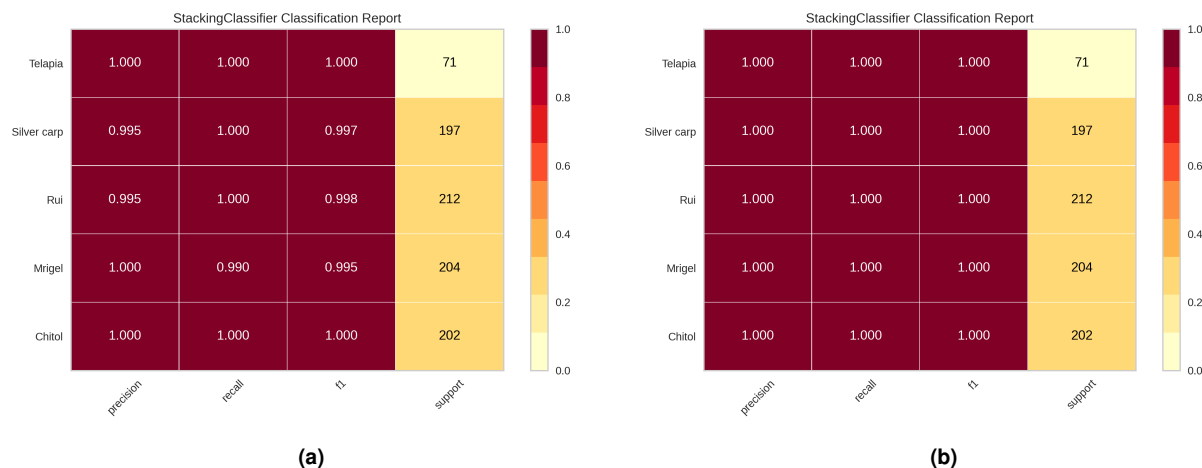


Fig. 4. (a) and (b) describe the classification report for model 1 and 2 respectively

and 2. That produces a good metric described in fig. 4.

Table 5. Comparison of the accuracy of various algorithms

Author Name	FC Algorithm Name	Accuracy
KAYA [17]	ANN	98.88%
Alsmadi [4]	HGAGD-BPC	96%
Hnin [14]	SVM	100%
Qin [25]	linear SVM classifier	98.57%
Matai [20]	PCA algorithm	100%
Ali-Gombe [3]	Deep CNN	97.20%
Kutlu [19]	Nearest neighbour	99%
Taheri-Garavand [30]	Deep CNN	98.21%
Abinaya [2]	NBC and DLN	98.60%
This article	Proposed model	99.77% and 100%

### 4.3 Comparison Analysis

The performance of the proposed classifier is compared using other types of classifiers from different suggested models as in table 5. This table describes that the recent works produce a good accuracy but comparatively proposed model stands better than the others.

## 5 Conclusion

A fish classification technique with salient object detection has been proposed in this paper to overcome the background variant issue on FC.

It has several steps of preprocessing approaches like image resizing and background removal. Afterward, a feature descriptor layer is used.

As previously illustrated, that preprocessing technique already separated many features depending on shape and color gradient, the ensemble layers with SVM, KNN, Logistic Regression, and Decision Tree play a nice role in the classification of the fishes. For the reason above, the proposed methodology stands good with high accuracy.

Our dataset consists of five classes of fish. The tested result is 99.77% on model-1 and 100% on model-2. The final decision is made from those ensembles depending on the high accuracy. Finally, our test results are compared to the proposed techniques illustrated on [5] and it stands tremendously good amongst all results illustrated there.

Moreover, the future enhancements are as follows: (i) The addition of a transfer learning approach with HOG images for more robust feature selection and dimensionality reduction. So there can be added a DLN layer after the HOG feature selection layer to achieve. (ii) Morphometric analysis of actual fish from the fish images. As this paper used the shape feature for classification, the shape can also be used for the height, width, and weight comparison. Therefore generate an automatic system that can produce

those characteristics from an image. (iii) Dataset improvement. In the future, the dataset will have images of more than five classes of fish so that it can empower more to the proposed models.

## References

1. **Abeldaim, A. M., Houssein, E. H., Hassanien, A. E. (2018).** Color image segmentation of fishes with complex background in water. *International Conference on Advanced Machine Learning Technologies and Applications*, Springer, pp. 634–643.
2. **Abinaya, N., Susan, D., Kumar, R. (2021).** Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. *Ecological Informatics*, Vol. 61, pp. 101248.
3. **Ali-Gombe, A., Elyan, E., Jayne, C. (2017).** Fish classification in context of noisy images. *International Conference on Engineering Applications of Neural Networks*, Springer, pp. 216–226.
4. **Alsmadi, M., Omar, K., Noah, S., Almarashdeh, I. (2011).** A hybrid memetic algorithm with back-propagation classifier for fish classification based on robust features extraction from PLGF and shape measurements. *Information Technology Journal*, Vol. 10, No. 5, pp. 944–954.
5. **Alsmadi, M. K., Almarashdeh, I. (2020).** A survey on fish classification techniques. *Journal of King Saud University-Computer and Information Sciences*.
6. **Baloch, A., Ali, M., Gul, F., Basir, S., Afzal, I. (2017).** Fish image segmentation algorithm (FISA) for improving the performance of image retrieval system. *International Journal of Advanced Computer Science and Applications(IJACSA)*, Vol. 8, No. 12.
7. **Chakraborty, S. C. (2021).** Fish.
8. **Chen, G., Sun, P., Shang, Y. (2017).** Automatic fish classification system using deep learning. *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, pp. 24–29.
9. **Chhabra, H. S., Srivastava, A. K., Nijhawan, R. (2020).** A hybrid deep learning approach for automatic fish classification. In *Proceedings of ICETIT 2019*. Springer, pp. 427–436.
10. **Dalal, N., Triggs, B. (2005).** Histograms of oriented gradients for human detection. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, IEEE, pp. 886–893.
11. **dos Santos, A. A., Goncalves, W. N. (2019).** Improving pantanal fish species recognition through taxonomic ranks in convolutional neural networks. *Ecological Informatics*, Vol. 53, pp. 100977.
12. **Fouad, M. M. M., Zawbaa, H. M., El-Bendary, N., Hassanien, A. E. (2013).** Automatic Nile tilapia fish classification approach using machine learning techniques. *13th international conference on hybrid intelligent systems (HIS 2013)*, IEEE, pp. 173–178.
13. **Ghose, B. (2014).** Fisheries and aquaculture in Bangladesh: Challenges and opportunities. *Annals of Aquaculture and Research*, Vol. 1, No. 1, pp. 1–5.
14. **Hnin, T. T., Lynn, K. T. (2016).** Fish classification based on robust features selection using machine learning techniques. In *Genetic and Evolutionary Computing*. Springer, pp. 237–245.
15. **Ibrahim, A., Ahmed, A., Hussein, S., Hassanien, A. E. (2018).** Fish image segmentation using salp swarm algorithm. *International Conference on advanced machine learning technologies and applications*, Springer, pp. 42–51.
16. **Jalal, A., Salman, A., Mian, A., Shortis, M., Shafait, F. (2020).** Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecological Informatics*, Vol. 57, pp. 101088.
17. **Kaya, E., Saritaş, İ., Taşdemir, Ş. (2017).** Classification of three different fish species by artificial neural networks using shape, color and texture properties. *7th International Conference on Advanced Technologies*.
18. **Kratzert, F., Mader, H. (2018).** Fish species classification in underwater video monitoring using convolutional neural networks. *EarthArXiv*.
19. **Kutlu, Y., Iscimen, B., Turan, C. (2017).** Multi-stage fish classification system using morphometry. *Fresenius Environmental Bulletin*, Vol. 26, No. 3, pp. 1911–1917.
20. **Matai, J., Kastner, R., Cutter, G., Demer, D. (2010).** Automated techniques for detection and recognition of fishes using computer vision algorithms. *NOAA Technical Memorandum NMFS-F/SPO-121*, Report of the National Marine Fisheries Service Automated Image Processing Workshop,

Williams K., Rooper C., Harms J., Eds., Seattle, Washington (September 4–7 2010).

21. **Miyazono, T., Saitoh, T. (2018)**. Fish species recognition based on CNN using annotated image. In *IT Convergence and Security 2017*. Springer, pp. 156–163.
22. **Mokti, M. N., Salam, R. A. (2008)**. Hybrid of mean-shift and median-cut algorithm for fish segmentation. *International Conference on Electronic Design, IEEE*, pp. 1–5.
23. **Pettersen, R., Braa, H. L., Gawel, B. A., Letnes, P. A., Sæther, K., Aas, L. M. S. (2019)**. Detection and classification of lepeophterius salmonis (Krøyer, 1837) using underwater hyperspectral imaging. *Aquacultural Engineering*, Vol. 87, pp. 102025.
24. **Pornpanomchai, C., Lursthut, B., Leerasakultham, P., Kitiyanan, W. (2013)**. Shape-and texture-based fish image recognition system. *Agriculture and Natural Resources*, Vol. 47, No. 4, pp. 624–634.
25. **Qin, H., Li, X., Liang, J., Peng, Y., Zhang, C. (2016)**. DeepFish: Accurate underwater live fish recognition with a deep architecture. *Neurocomputing*, Vol. 187, pp. 49–58.
26. **Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., Jagersand, M. (2020)**. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition*, Vol. 106, pp. 107404.
27. **Rekha, B., Srinivasan, G., Reddy, S. K., Kakwani, D., Bhattad, N. (2019)**. Fish detection and classification using convolutional neural networks. *International Conference On Computational Vision and Bio Inspired Computing*, Springer, pp. 1221–1231.
28. **Sayed, G. I., Hassanien, A. E., Gamal, A., Ella, H. A. (2018)**. An automated fish species identification system based on crow search algorithm. *International Conference on Advanced Machine Learning Technologies and Applications*, Springer, pp. 112–123.
29. **Sharmin, I., Islam, N. F., Jahan, I., Joye, T. A., Rahman, M. R., Habib, M. T. (2019)**. Machine vision based local fish recognition. *SN Applied Sciences*, Vol. 1, No. 12, pp. 1–12.
30. **Taheri-Garavand, A., Nasiri, A., Banan, A., Zhang, Y.-D. (2020)**. Smart deep learning-based approach for non-destructive freshness diagnosis of common carp fish. *Journal of Food Engineering*, Vol. 278, pp. 109930.

*Article received on 29/07/2021; accepted on 30/09/2021.  
Corresponding author is Monowar Hossain.*