SN Computer Science (2021) 2:273
https://doi.org/10.1007/s42979-021-00614-8

ORIGINAL RESEARCH

FishResNet: Automatic Fish Classification Approach in Underwater Scenario

Monika Mathur[1] · Nidhi Goel[1]

Received: 14 September 2020 / Accepted: 23 March 2021 / Published online: 14 May 2021

Abstract
Fish species classification in underwater images is an emerging research area for scientists and researchers in the field of image processing. Fish species classification in underwater images is an important task for fish survey i.e. to audit ecological balance, monitoring fish population and preserving endangered species. But the phenomenon of light scattering and absorption in ocean water leads to hazy, dull and low contrast images making fish classification a tedious and tough task. Convolutional Neural Networks (CNNs) can be the solution for fish species classification problem but the scarcity of ample fish images leads to the serious issue of training a neural network from scratch. To overcome the issue of limited dataset the present paper proposes a transfer learning based fish species classification method for underwater images. ResNet-50 network has been used for transfer learning as it reduces the vanishing gradient problem to minimum by using residual blocks and thus improving the accuracies. Training only last few layers of ResNet-50 network with transfer learning increases the classification accuracy despite of scarce dataset. The proposed method has been tested on two datasets comprising of 27, 370 (i.e. large dataset) and 600 images (i.e. small dataset) without any data augmentation. Experimental results depict that the proposed network achieves a validation accuracy of 98.44% for large dataset and 84.92% for smaller dataset. With the performance analysis, it is observed that this transfer learning based approach led to better results by providing high precision, recall and F1score values of 0.94, 0.85 and 0.89, respectively.

Keywords Convolutional Neural Network · Fish species · Transfer learning · Underwater images · ResNet-50 · AlexNet

Introduction

Object recognition and classification is highly researched problem in underwater image and video analysis [1–3]. There is fast development in ocean observation in the last decade but this is quite challenging task due to underwater environment which involves scattering and absorption of light [4]. These properties of light in water leads to light attenuation causing reduced visibility and thus, underwater videos and images have usually low quality. This makes fish species classification as the most difficult tasks for researchers, biologists and marine scientists. Ocean researchers are interested in using underwater videos and images for fish population analysis, fish species classification and size measurement etc. [5–7].

Underwater fish images required for these researches are usually collected by scuba divers or autonomous underwater vehicles (AUVs) and then annotated physically by marine scientists and biologists. These processes undoubtedly consume lot of time, resources and manpower. But now a days automatic systems implementing deep learning are more common for challenging problem of fish species classification with scarce and low quality data. These automatic systems can classify the fish species more accurately and without human intervention. Earlier analysis and classification of fish species is done manually that involves destructive and time consuming methods like capturing of fishes or visual census by deep sea divers.

This paper aims at finding a solution to underwater fish classification with high accuracy. Deep neural networks can be the solution to above mentioned problem. But learning a deep neural network from scratch requires a huge dataset [8, 9]. Datasets available for fish classification are of small

✉ Nidhi Goel
nidhi.iitr1@gmail.com

Monika Mathur
monika009phd0215@igdtuw.ac.in

[1] Department of ECE, IGDTUW, Kashmere Gate, Delhi-06, India

size which are not sufficient for learning. So, an automatic system which uses transfer learning for fish classification is proposed in this paper. Transfer learning based methods are fast and requires less data for training as they uses only last few layers of pre-trained network for classification. Further, the proposed classification method will be helpful in monitoring the fish biodiversity, their population and distribution in aquatic ecosystem.

The proposed paper is organized in the following sequence: Sect. "Literature Survey" focuses on the literature survey and Sect. "Proposed Methodology" elaborates the proposed methods along with description of transfer learning. Section "Results" demonstrates the result and comparative analysis. Conclusion of paper is presented in Sect. "Conclusion".

Literature Survey

Fish species classification is a tough job because of the challenging underwater environment. Many researchers and scholars are working in this domain since last decade and various novel approaches have been proposed for fish species classification.

Cabreira et al. [10] proposed a technique for the automatic classification of fish species depending on their echo recordings. Three parameters namely bathymetric, energetic and morphometric were mined from the echo recordings to serve as inputs to the Artificial Neural Networks (ANNs) used for testing. Classification accuracy of 96% was obtained, depending on the type of school parameters and networks utilized. Spampinato et al. [11] proposed an automatic fish classification method by combining artificial immune systems and image analysis techniques. This automatic framework uses combination of Scale-Invariant Feature Transform (SIFT) and Principal Component Analysis (PCA) for identifying shape, appearance and motion of the fishes. Adaptive Radius Immune technique and Artificial Immune Network are used for clustering the identical species; classification accuracy of 92% is achieved using nearest neighbour algorithm [11].

Fish species classification method proposed by Castillo et al. [12] uses multiclass support vector machine (SVM) and two supervised ANN namely; Probabilistic Neural Network (PNN) and Multilayer Perceptron (MLP). Acoustic records were used as descriptor for the neural networks. Obtained results shows that both MLP and SVM methods with accuracy 89.5% outperforms the PNN with classification accuracy of 79.4%. Boom et al. [13] proposed a heuristics tree based classification method and finally compares it to a state-of-the-art tree on a live fish image dataset. It outperforms the baseline method by achieving an accuracy of 90.0%. Hu et al. [14] proposed a novel approach by utilising the multiclass support vector machine along with texture and color features for fish species categorization.

Zawbaa et al. [15] proposed an automatic fish classification technique based on SVM classifier and Speeded Up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT) algorithms for feature extraction. Result analysis shows that the SVM based classification algorithm outperformed state of art techniques like ANNs and k-nearest neighbour (k-NN). A hierarchical categorization method for fish species recognition is proposed by Kuo et al. [16]. All levels of the species hierarchy undergoes partial categorization for coarse-to-fine cataloguing; discriminative feature descriptors are generated from species anatomical parts. Experimentation proves that the partial classification algorithm works well on imbalanced dataset by achieving a classification accuracy of 94% and partial decision rate of 5%.

Ogunlana et al. [17] also proposed a SVM based technique for fish classification. The technique uses the five fin parameters (caudal, pelvic, dorsal, pectoral and anal) and the fish body for shape feature extraction. Results shows a classification accuracy of 78.59%, which outperformed K-mean clustering, K-NN and ANN algorithms. Shang et al. [18] proposed Deep-CNN approach on the Fish for knowledge dataset consisting of 27, 370 fish images and achieved an accuracy of 98.57% in fish classification. Rodrigues et al. [19] proposed five different schemes for the species recognition based on combination of clustering algorithms (aiNet, k-means and ARIA), input classifiers (k-means, SIFT and k-NN) and feature extraction methods (PCA+SIFT+VLAD, SIFT, PCA). These schemes uses two datasets i.e. images of four species in natural surroundings and six species preserved in formaldehyde solution. Results show that these schemes are less time consuming and cheap.

Salman et al. [20] proposed a CNN which uses species–dependent hierarchical features and evades the requirement of extracting them from unprocessed fish images. The CNN based method also performs good for test images which does not belongs to the training data. LifeCLEF15 and LifeCLEF14 fish datasets were used for results analysis and a classification accuracy of 90% is achieved. Liang et al. [21] also proposed deep learning based method for feature extraction of fish images. Two convolutional layers of neural network use Principal Component Analysis (PCA). Further, binary hashing and block wise histogram are used for non-linear layer and the feature pooling layer, respectively. Finally, classification is done using linear SVM classifier. Automated fish classification method based on image processing, deep learning and convolutional neural networks was presented by Rathi et al. [22] with an accuracy of 96.29%.

Demertzis et al. [23] proposed an automatic Machine Hearing Framework (MHF) for Marine Species classification and recognition in underwater environment. It identifies fish species based on their sounds. Online MIGRATE ELM Autoencoder has been used for audio recognition. An automated system based on CNN was proposed by Iqbal et al. [24] for fish species identification. It uses reduced form of AlexNet and utilizes only four convolutional and two fully connected layers instead of five convolutional and three fully connected layers. Obtained accuracy of 90.48% proves that Alexnet with less number of layers is efficient even on testing data.

Chuang et al. [25] proposed an approach which extracts fish features by unsupervised learning and finally, clusters species using an error resilient classifier. Further, Qiu et al. [26] presents a novel method for fine-grained fish species recognition by transfer learning and squeeze-and-excitation networks. It utilizes data augmentation and super-resolution reconstruction techniques for improving and enhancing small scale and low quality image dataset.

Hridayami et al. [27] proposed a transfer learning model based on VGG16 for fish species classification. The dataset for classification consists of 50 species and is divided into four different types i.e. canny filter image, RGB color space image, blending image + RGB image and simple blending image. The result analysis indicates that blending image + RGB image achieves the Genuine Acceptance Rate (GAR) of 96.4%.

Most of the above mentioned papers either use the large datasets to train their network from scratch or uses data augmentation techniques to increase the number of training images to enhance the accuracies artificially. But the proposed method uses the transfer learning-based approach to achieve better accuracies even with limited or small dataset. Next section describes the proposed method along with describing the transfer learning.

Proposed Methodology

Fish species classification and recognition is a challenging and tough task due to underwater environment. Challenges hindering correct fish recognition comprises of noise, distortion, overlap, segmentation error, and obstructions. Various techniques including K-mean clustering, K-Nearest neighbour (KNN) and neural network, are generally employed to solve these challenges but with certain limitations, which bounds the classification accuracy. Most of the above mentioned methods are computationally complex and requires huge data for accurate classification. To overcome these limitations and to achieve high accuracy with limited and untrained dataset, transfer learning based approach is proposed in this paper.

Transfer learning is a machine learning technique which builds accurate deep learning models in a timesaving way [28]. In transfer learning, learning process is not done from scratch; rather it extracts the features from pretrained layers of the network while solving a different problem. In this way transfer learning based networks make use of earlier learnings of network and resist starting it from scratch. Transfer learning generally makes use of pre-trained networks. Pre-trained networks are the networks that are trained on a large dataset consisting of millions of images to solve a problem identical to the one that we are interested in. Instead of training a network (deep networks) from scratch which is computationally complex, time consuming and costly, the proposed work uses pretrained networks (e.g. AlexNet, VGG, MobileNet, GoogleNet, ResNet etc.) to save both time and resources.

Canziani et al. [29] presents a comprehensive review of pretrained networks' performance using images from the ImageNet challenge [30]. The present paper proposes transfer learning for fish species classification. The major cause of overfitting and loss of generality in the neural networks is due to lack of high number of data points for the model to learn from. The use of transfer learning compensates for this as the existing neural networks are already trained on millions of images. These models when fine-tuned for a particular application, for example classification of fish species in this case, decrease the variance and bias for the model.

Selection of CNN for Transfer Learning

Selection of pre-trained network is a tough task as number of architectures are available in the literature with good performances. Table 1 shows the top1 and top5 accuracy for various CNNs [31]. Top1 accuracy is the conventional version of accuracy, it only considers top one class with the highest probability, whereas Top-5 accuracy uses top five classes instead of 1. On basis of best top1 and top5 accuracy shown in Table 1, the present paper uses Resnet50 as pre-trained network for transfer learning. Resnets are a kind of CNNs called Residual Networks. They are very deep compared to Alexnet and VGG, and Resnet 50 refers to a 50 layers Resnet. They overcome the problems of vanishing gradients which is main issue in other deep networks with more layers.

Training deep neural networks (consisting of large layers) diminishes the gradient dramatically because

Table 1 Comparison of top1 and top5 accuracies of various pre-trained neural networks

Architecture	Top1 accuracy (%)	Top5 accuracy (%)
AlexNet	57.1	80.2
GoogleNet	69.8	89.3
VGG	70.5	91.2
ResNet-50	75.2	93

it propagates backward through the network and error reduces to nearly zero as it reaches initial layers of the network. This problem is known as the problem of vanishing gradients. Vanishing gradients distract the parameters and it is tough to find in which direction they should move in order to increase or decrease the cost function. To overcome this problem of vanishing gradients and thus improving the accuracy, Resnet networks utilizes residual connections between its layers, i.e. the output of a layer is a convolution of its input and output [32]. This is shown in Fig. 1.

The basic idea is to add an identify connection every few layers that adds the source of the block, x to the output of the block $f(x)$, resulting in the final output of $H(x) = f(x) + x$. The name residual networks comes from the fact that network is learning $f(x) = H(x) - x$, the residual when input is subtracted from output.

Use of normal distribution to initiate the weights in other networks cannot effectively solve the problem of vanishing gradients and thus decrease their performances. To overcome the problem of vanishing gradients, residual network ResNet-50 is used for fish species classification utilizing the concept of transfer learning i.e. it take layers from ResNet-50 trained on a large data set and fine-tune them on a given fish data set. This fine-tuned network is named as FishResNet.

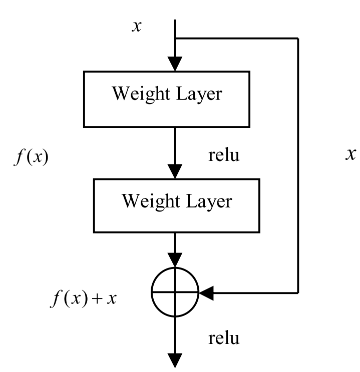


Fig. 1 Concept of residual connections

FishResNet Network

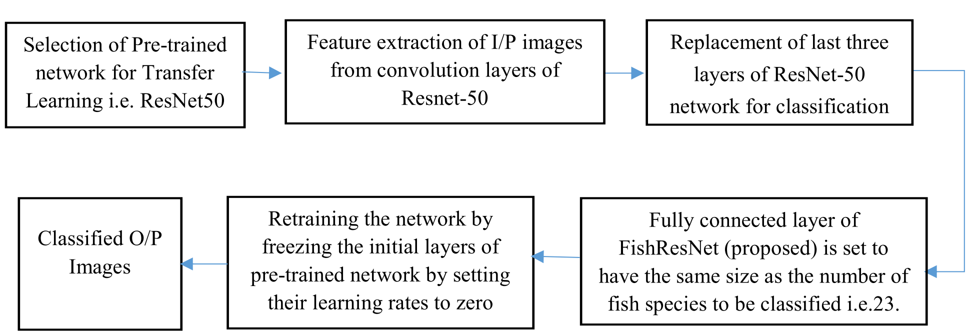
The last few layers of the neural network have been modified to apply transfer learning to it. The pre-trained neural networks are already trained to learn rich feature representations for various types of images. In present paper pre-trained network, ResNet-50 has been used for transfer learning. ResNet-50 uses input images of size $224 * 224 * 3$, with 3 as number of color channels. Flowchart for the proposed method is shown in the Fig. 2

Image features have been extracted using the convolutional layers of ResNet-50 network but classification has been done using the final softmax and classification layers. Fish classification in FishResNet has been done by replacing the last three layers of ResNet-50 network and retraining it with these modified layers. These three layers, fc1000, fc1000-softmax and classification layer-fc1000 shown in Fig. 3 contain information about combining the image features extracted from earlier layers into class probabilities and lables. The FishResNet replaces these last three layers of ResNet-50 network by a fully connected layer, a softmax layer, and a classification output layer as shown in Fig. 3.

The final fully connected layer of FishResNet is set to have the same size as the number of fish species to be classified i.e. 23. The learning rate of the fully connected layer in FishResNet is increased to learn faster in new layers. To complete the connections in FishResNet, the last transferred layer i.e. avg-pool left in ResNet-50 is connected to these new layers (fully connected, softmax, and classification output) of FishResNet. These connections of ResNet-50 and transferred ResNet-50 (i.e. FishResNet) are shown in Figs. 3 and 4 respectively.

Initial Layers of FishResNet network are freezed by setting their learning rates to zero. During network training, gradients of these freezed layers are not computed because of zero learning rates and this will considerably improves the training speed. Freezing of layers also prevent overfitting problem in case of small datasets. The FishResNet has same number of layers as ResNet-50, but the the earlier layers are freezed or their learning rate has been set to zero. FishResNet uses tanh as activation function, softmax as classifier and Adamax as optimizer. Leaning rate for last

Fig. 2 Flowchart for the proposed FishResNet Network



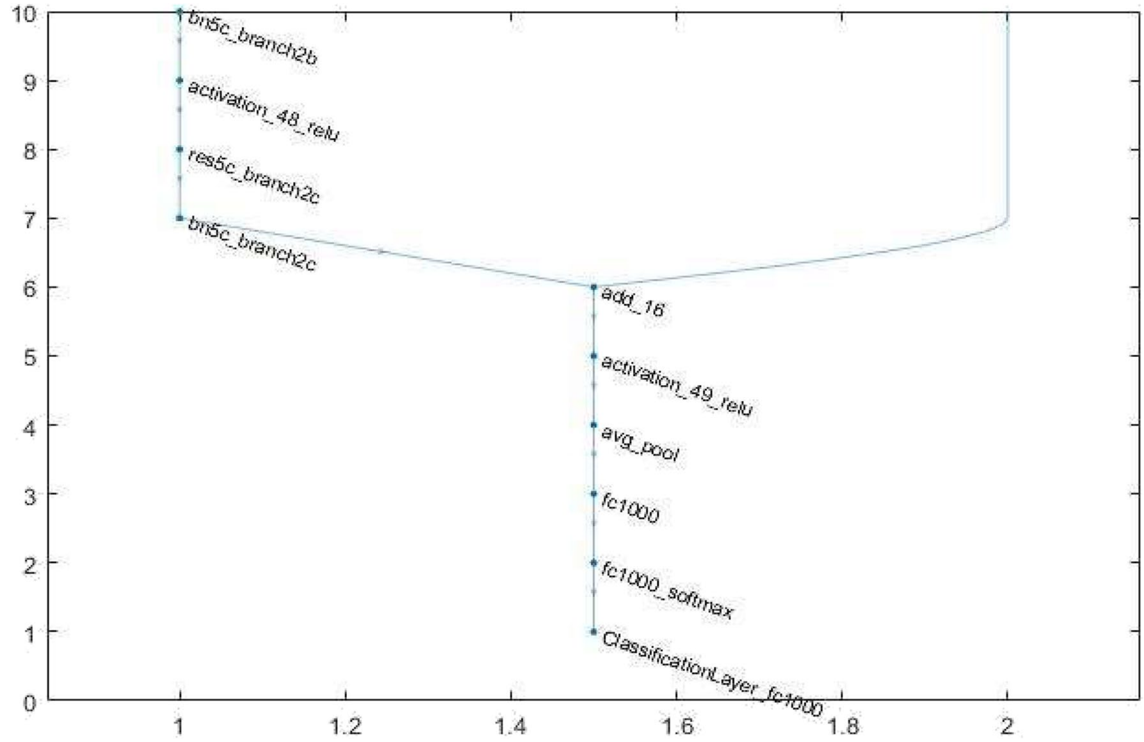


Fig. 3 Original connection of last three layers (i.e. fc1000, fc1000-softmax and classification layer-fc1000) in ResNet50 network

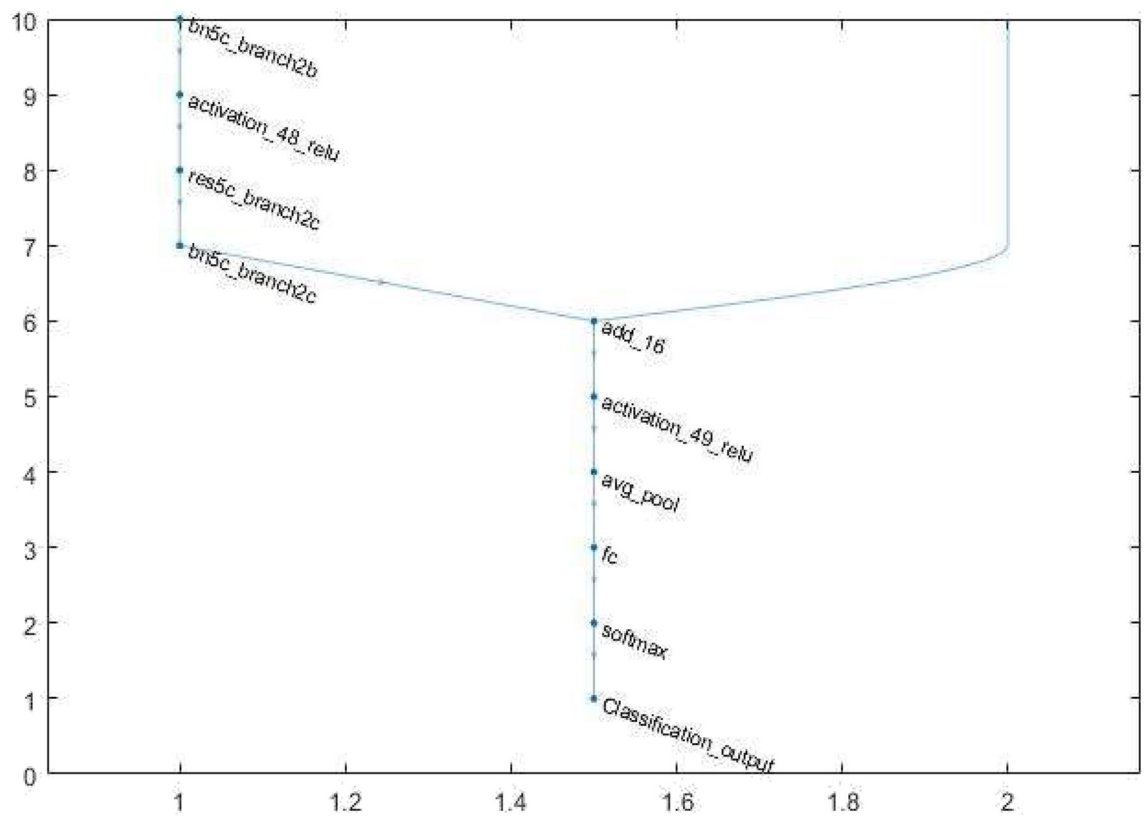


Fig. 4 Connection of last three layers (i.e. fully connected, softmax, and classification output) in fine tuned ResNet50 Network i.e. FishResNet

three layers is set to 0.001 and each run cycle comprises of 5 epochs with batch size of 64. 70:30 ratio has been used for training and test images. FishResNet is evaluated using different optimizers with Tanh as activation function and softmax as classifier. Table 2 below compares the accuracies with different optimization techniques based on F4k dataset. Results in Table 2 shows that Adamax is the best optimizer of our problem statement. So, FishResNet classifies the fish species by transfer learning or fine tuning the ResNet-50 network based on above parameters.

Table 2 Comparison of FishResNet accuracies with different optimizers based on F4k dataset

Model	Activation function	Optimizer	Accuracies (%)
Model A	tanh	adam	90.27
Model B	tanh	rmsprop	92.71
Model C	tanh	adadelta	94.49
Model D	tanh	adamax	98.44

Results

Datasets

Effectiveness of the proposed networks is evaluated on Fish Recognition dataset made by the Fish4Knowledge (F4K) project [33] and QUT dataset [34]. The fish dataset, F4K used for training is acquired from a live underwater video dataset. This data is divided into 23 fish species. These species are physically categorized by marine biologists. Out of 27, 370 fish images 70% (19, 159 images) have been used as training dataset, whereas remaining 30% (8, 211 images) have been used as validating dataset. Matlab 2018a and NVIDIA Quadro K2200 GPU are used for implementation of algorithms and evaluation of results.

The QUT fish dataset comprises of 3960 images collected from 468 species in undercontrolled, out-of-the-water and underwater conditions. For proposed methods 37 species have been chosen for classification comprising of 600 images in underwater condition. 70 : 30 ratio has been used for training and test images. Figure 5 shows the 23 fish species to be classified whereas Table 3 represents the 23 fish species and number of fishes in each category. The proposed methods have been tested on these large (F4K) as well as small (QUT) datasets.

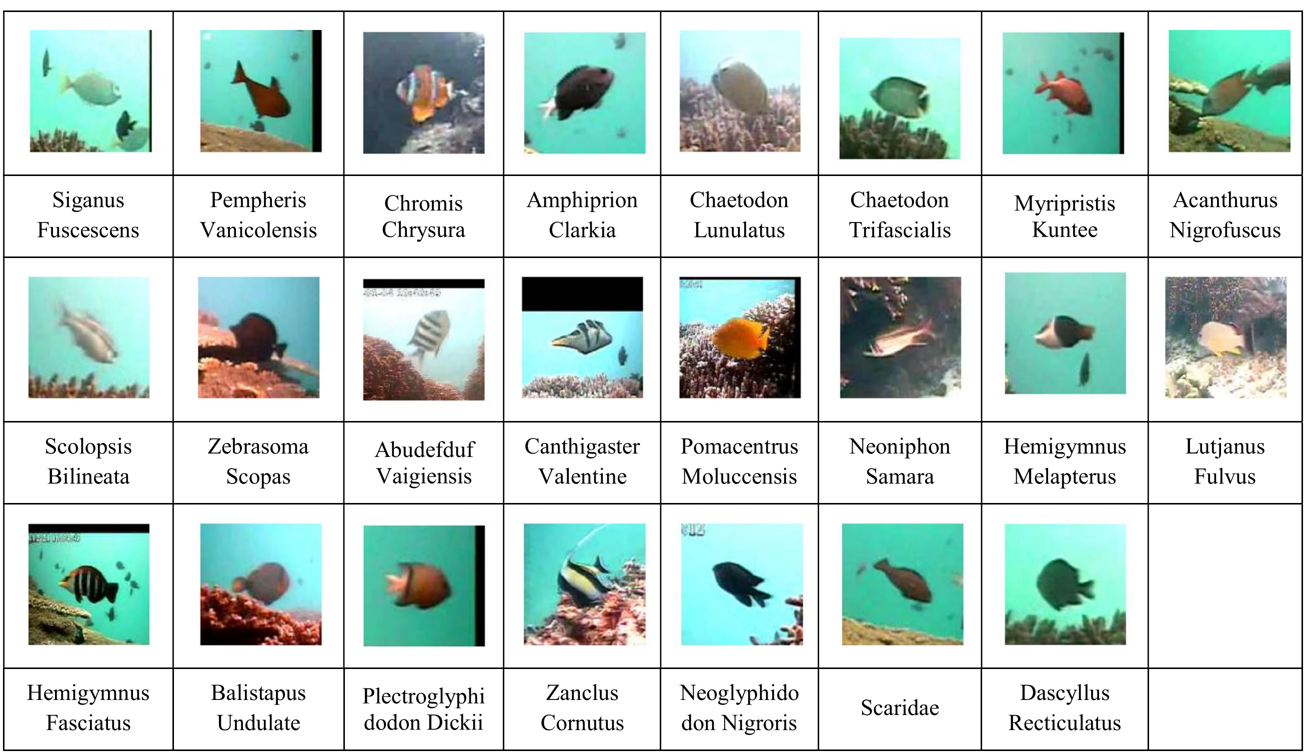


Fig. 5 Examples of 23 fish species from F4K dataset

Table 3 F4K dataset with 23 fish species and number of fishes in each category

Fish species	No. of samples	Fish species	No. of samples
Amphiprion Clarkia	4049	Dascyllus Recticulatus	12,112
Chaetodon Lunulatus	2534	Chromis Chrysura	3593
Myripristis Kuntee	450	Plectroglyphidodo Dickii	2683
Neoniphon Samara	299	Lutjanus Fulvus	206
Acanthurus Nigrofuscus	218	Hemigymnus Fasciatus	241
Chaetodon Trifascialis	190	Canthigaster Valentine	147
Pomacentrus Moluccensis	181	Scolopsis Bilineata	49
Abudefduf Vaigiensis	98	Zebrasoma Scopas	90
Zanclus Cornutus	21	Scaridae	56
Balistapus Undulates	41	Hemigymnus Melapterus	42
Siganus Fuscescens	25	Pempheris Vanicolensis	29
Neoglyphidodon Nigroris	16		

Result Evaluation and Comparisons

The proposed FishResNet method have been evaluated using F4K and QUT datasets. FishResNet achieves an test accuracy of 98.44% on large dataset for 5 epochs in 1387 min and 5 s with 0.001 learning rate. Training and Loss curves of FishResNet method for large dataset are shown in Fig. 6. Blue line in the upper part of Fig. 6 shows the training accuracy whereas the bold blue line represents the smoothed training accuracy. Validation accuracy is shown by the black dotted line.

Minimal gaps between training and validation accuracies indicates the good fit for the curve and hence high performance. Lower part of Fig. 6 shows the loss curve corresponding to large datatset. Red line represents the training loss whereas the bold red line shows the smoothed training loss. Validation loss is represented by black dotted line. Loss curves show a good fit as validation and training loss values are decreasing to a stability point with a minimum gap between these loss values. Low loss values indicate the good performance of the FishResNet Network.

Similarly, the results have also been evaluated for small dataset to measure the changes in the accuracy and speed. As the data is small, both the accuracy and time consumed is reduced for the proposed method. FishResNet achieves a test accuracy of 84.17% for 5 epochs and 0.001 learning rate in 14 min and 45 s. Training and Loss curves of FishResNet method for small dataset are shown in Fig. 7. Blue line in the upper part of Fig. 7 shows the training accuracy whereas the bold blue line represents the smoothed training accuracy. Validation accuracy is shown by the black dotted line.

The curves of Fig. 7 shows overfitting results because the increasing gap between the training and validation accuracies does not accurately estimates the response on new images, which were not present in the training dataset. Lower part of Fig. 7 shows the loss curve for large datatset. Red line represents the training loss whereas the bold red line shows the smoothed training loss. Validation loss is represented by black dotted line. Training loss curve decreases continuously as model learns whereas the validation loss plot continues to decrease upto certain point and then increases again showing overfitted results. The results for both small and large datasets are tabulated in Table 4.

Results in Table 4 depict that performance of FishResNet is better for large dataset both in terms of accuracy and time consumption. The results for the small dataset are overfitted as seen from the loss and accuracy graphs of Fig. 7. This is due to the fact that exceptions in the training data (i.e. statistical noise or random fluctuations) can be easily remembered by the model with less data. It leads to high accuracy on training but low accuracy on test set as the model generalizes what it has learned from the small training set.

The performance of proposed model has been further tested on the basis of Precision, recall and F1 score, which are generally used in classification problems with imbalanced datasets. These scores are calculated on the F4K dataset. F1 score is more realistic measure of a test's performance by providing a balance between precision and recall. The precision defined in Eq. 1 is ratio of correctly classified images to the total number of predicted images which lies in a particular class. On the other hand, Recall is ratio of correctly classified images to total number of images which belongs to a particular class and is defined in Eq. 2.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Positives}}, \tag{1}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Negatives}}, \tag{2}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{3}$$

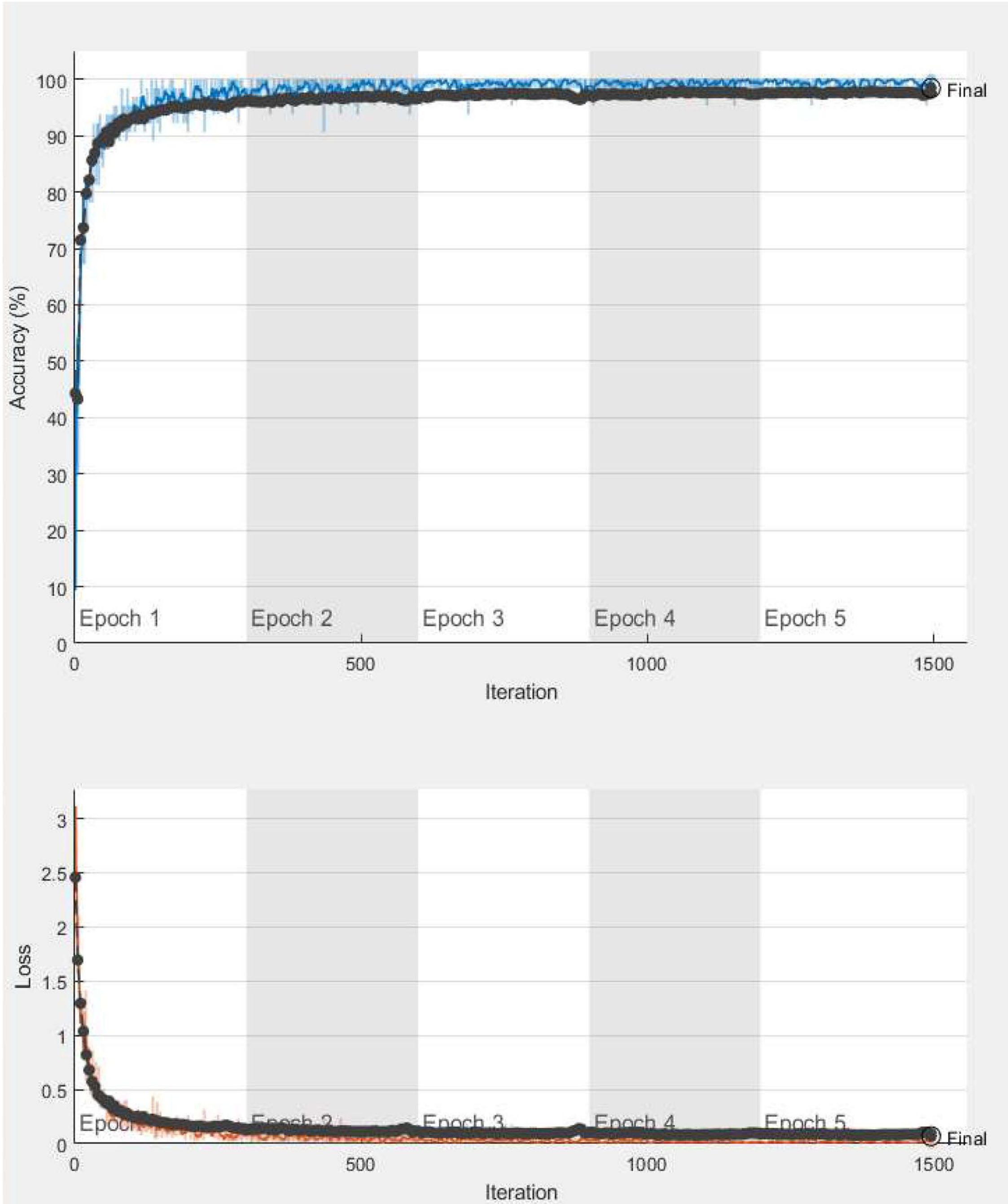


Fig. 6 Training and loss curves of FishResNet network on F4K dataset

F1 score defined in Eq. 3 is harmonic mean of Precision and Recall values. Precision, Recall and F1 scores for 23 species used in classifications are shown in Table 5. It can be observed that Precision values are higher than recall values in most of the cases, indicating that false negatives are more common than false positives in the proposed model. High value of precision shows that high proportion of predicted positives is truly positive indicating that the number of false positives is very low and FishResNet model is correctly classifying the species. It can be observed from Table 5 that macro-averaged Precision, Recall and F1-score for proposed model are 0.94, 0.85 and 0.89 respectively, which indicates that performance of proposed model is quite good as scores are approaching 1 i.e. maximum value for these measures.

The results of 23 fish species classification using FishResNet are shown in Table 6 in terms of accuracy

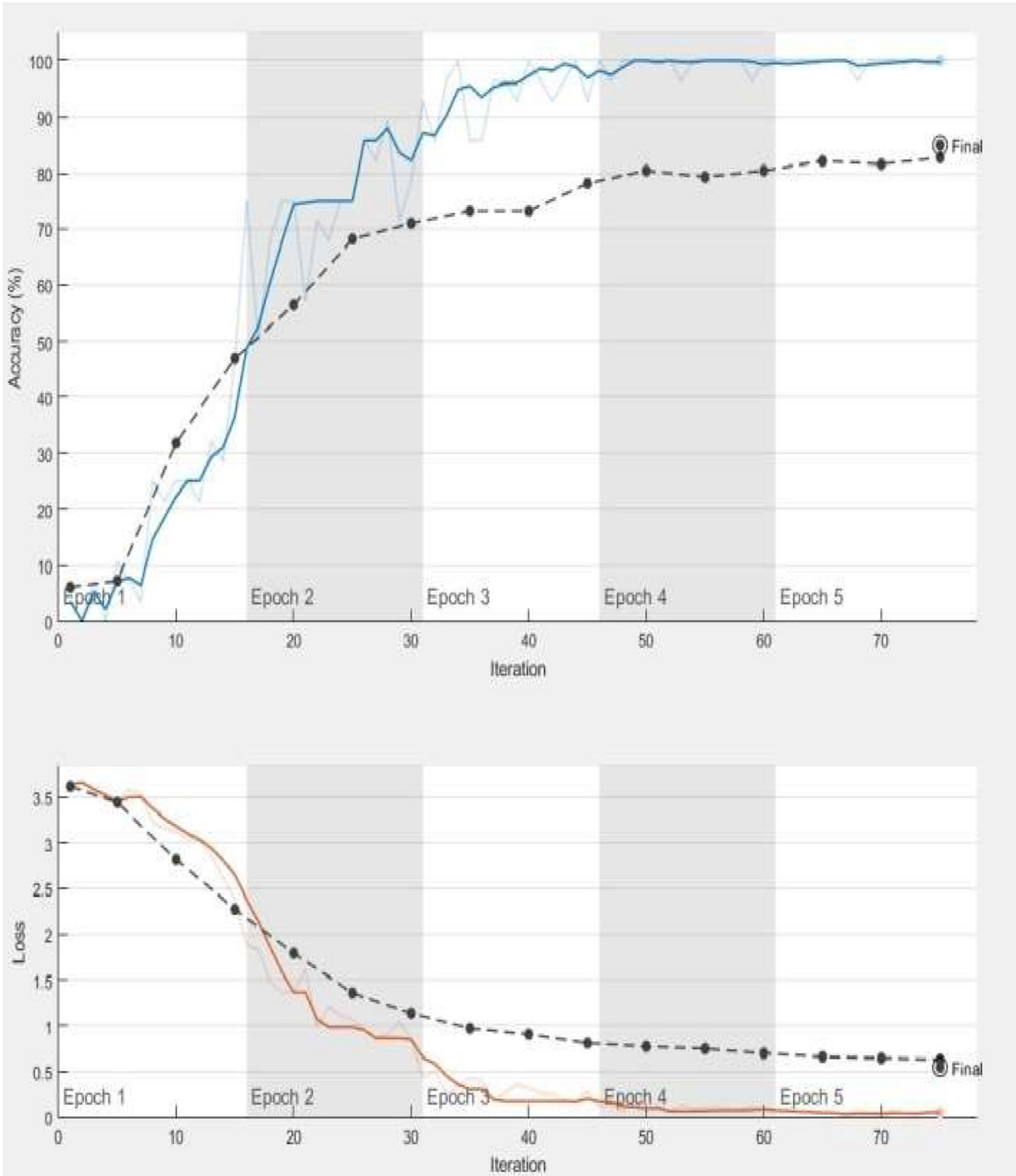


Fig. 7 Training and loss curves of FishResNet network on QUT dataset

Table 4 Results of FishResNet in terms of accuracy and time consumption

	FishResNet	
	Accuracy (%)	Time
QUT dataset	84.92	14 min 45 s
F4K dataset	98.44	1387 min 05 s

along with its comparison to existing methods on the F4k dataset. Various machine learning tools and techniques have been used as reference methods for comparisons. Methods based on LDA as feature extractor and SVM as classifier [35] obtained an accuracy of 80.14% by using fish images without background. Validation accuracy of 89.79% has been achieved by nearest neighbour method for fish species classification. A simple SVM classifier based approach obtains an accuracy of 82.98% by training the model on the raw pixels.

Table 5 Precison, recall and F1 score for 23 fish species on F4K dataset

Species	Precision	Recall	F1 Score	Species	Precision	Recall	F1 Score
Amphiprion Clarkia	1	1	1	Neoniphon Samara	0.97	0.78	0.86
Abudefduf Vaigiensis	0.86	0.66	0.74	Myripristis Kuntee	1	1	1
Canthigaster Valentine	0.89	0.73	0.81	Plectroglyphidodon Dickii	1	1	1
Acanthurus Nigrofuscus	1	0.82	0.90	Neoglyphidodon Nigroris	1	0.86	0.92
Hemigymnus Fasciatus	0.34	0.86	0.49	Zanclus Cornutus	0.95	0.56	0.70
Balistapus Undulates	1	0.69	0.82	Pempheris Vanicolensis	1	1	1
Chaetodon Lunulatus	1	1	1	Pomacentrus Moluccensis	0.98	0.93	0.95
Chaetodon Trifascialis	0.92	0.83	0.87	Scaridae	1	1	1
Chromis Chrysura	1	0.66	0.80	Scolopsis Bilineata	1	0.97	0.98
Dascyllus Recticulatus	1	1	1	Lutjanus Fulvus	1	0.98	0.99
Zebrasoma Scopas	1	0.9	0.94	Hemigymnus Melapterus	0.87	0.60	0.71
Siganus Fuscescens	1	0.77	0.8	Macro-averaged Scores for 23 fish species	0.94	0.85	0.89

Further, softmax classifier [21] in DeepFish model achieves an accuracy of 87.56% whereas, VLFeat method [36] used for comparison gives an accuracy of 93.58%. Results from DeepFish architecture [21] with data augmentation gives a test accuracy of 98.23%. The compaison results are listed in Table 6. Deep-CNN [18], uses deep learning architecture for designing its three convolutional layers. It is trained from scratch for fish and plankton classification. It achieves an accuracy of 98.57%. Alex-FT-Soft [37] uses AlexNet for feature extraction, whereas softmax classifiers for fish recognition and gives an accuracy of 96.61%.

The high values of validation or test accuracies demonstrate that proposed networks with transfer learning outperform the existing methods for fish species classification. Better performance of FishResNet is due to the use of ResNet-50 network as it uses batch normalization at its core and adjusts the input layer to increase the performance of the network. Batch normalization also mitigates the problem of covariate shift. Use of the identity connection in ResNet-50 network protects it from vanishing gradient problem and bottleneck residual block design increase the performance of the network. It further indicates that transfer learning based neural networks are much faster as there is no need to train them for as many epochs as a new model would require [38] i.e. to train from beginning. Transfer learning based networks also require comparatively less images as compared to huge datasets (million or more) required for training a model from ground level. Hence, proposed methods are accurate solution for fish categorization with minimum labour [39].

Table 6 Comparitive Analysis of FishResNet with various models on F4K dataset

Classification models	Accuracy (%)
Raw-Pixel Nearest Neighbour [35]	89.79
Raw-Pixel SVM [35]	82.92
LDA+SVM [35]	80.14
VLFeat Dense-SIFT [36]	93.58
Raw-Pixel Softmax [21]	87.56
DeepFish-SVM [21]	98.23
DeepFish-Softmax-Aug [21]	92.55
Alex-FT-Soft [37]	96.61
Deep-CNN [18]	98.57
FishResNet (Proposed)	98.44

Conclusion

Automatic Fish recognition and classification is important for marine biologists for studying the variety of fishes. Since traditional methods are not eco-friendly as they affect fish behaviour by physically capturing them and also demands time and labour costs but transfer learning based FishNet and FishResNet networks provides a cost-effective and efficient solution to it. In the scenarios, where availability of sufficient data is an issue; transfer learning proves to be a better solution for this type of problem as pre-trained networks are used with fine tuning according to desired requirements and as feature extractor. Experimental results demonstrate that the proposed method outperform various state-of-the-art methods used for fish species classification. The obtained results also reflect that the networks trained with transfer learning perform well than networks trained from scratch both in terms of computation complexity and accuracy. Further, future work would include the classification of multiple objects in the underwater images and videos instead of single object classification.

Declaration

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

References

1. Roberts PLD, Jaffe JS, Trivedi MM. Multiview, broadband acoustic classification of marine fish: a machine learning framework and comparative analysis. IEEE J Oceanic Eng. 2011;36(1):90–104.
2. Wang G, Hwang J, Williams K, Wallace F, Rose CS. Shrinking encoding with two-level codebook learning for fine-grained fish recognition. In: 2016 ICPR 2nd Workshop on Computer Vision for Analysis of Underwater Imagery (CVAUI). 2016; pages 31–36.
3. Li X, Shen H, Zhang L, Zhang H, Yuan Q, Yang G. Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning. IEEE Trans Geosci Remote Sens. 2014;52(11):7086–98.
4. Schettini R, Corchs S. Underwater image processing: state of the art of restoration and image enhancement methods. EURASIP J Adv Signal Process. 2010;2010(1):1–14.
5. Hasija S, Buragohain MJ, Indu S. Fish species classification using graph embedding discriminant analysis. In: 2017 International Conference on Machine Vision and Information Technology (CMVIT). 2017; pages 81–86.
6. Sbragaglia V, Marini S, Fanelli E. Tracking fish abundance by underwater image recognition. Sci Rep. 2018;8:1–12.
7. Andrialovanirina N, Ponton D, Behivoke F, Mahafina J, Léopold M. A powerful method for measuring fish size of small-scale fishery catches using imagej. Fisheries Res. 2020;223:1–8.
8. Kuanar S, Rao KR, Bilas M. Adaptive cu mode selection in hevc intra prediction: a deep learning approach. Circuits Syst Signal Process. 2019;38:5081–102.
9. Kuanar S, Rao KR, Mahapatra D, Bilas M. Night time haze and glow removal using deep dilated convolutional network, 2019. arXiv:1902.00855.
10. Cabreira A, Tripode M, Madirolas A. Artificial neural networks for fish-species identification. ICES J Mar Sci. 2009;66:1–11 (**06**).
11. Spampinato C, Giordano D, Di Salvo R, Jessica YH, Burger C, Fisher RB, Nadarajan G. Automatic fish classification for underwater species behavior understanding. In: ARTEMIS@ ACM Multimedia. 2010; pages 1–6.
12. Robotham H, Bosch P, Gutiérrez-Estrada JC, Castillo J, Pulido-Calvo I. Acoustic identification of small pelagic fish species in chile using support vector machines and neural networks. Fish Res. 2010;102(1):115–22.
13. Huang PX, Boom BJ, Fisher B. Hierarchical classification for live fish recognition, In: BMVC, 2012. pp 1–10.
14. Hu J, Li D, Duan Q, Han Y, Chen G, Si X. Fish species classification by color, texture and multi-class support vector machine using computer vision. Comput Electron Agric. 2012;88:133–40.
15. Fouad MMM, Zawbaa HM, El-Bendary N, Hassanien AE. Automatic nile tilapia fish classification approach using machine learning techniques. In: 13th International Conference on Hybrid Intelligent Systems (HIS 2013). 2013; pages 173–178.
16. Chuang M, Hwang J, Kuo F, Shan M, Williams K. Recognizing live fish species by hierarchical partial classification based on the exponential benefit. In. 2014 IEEE International Conference on Image Processing (ICIP). 2014; pages 5232–5236.
17. Ogunlana SO, Olabode O, Oluwadare SAA, Iwasokun GB. Fish classification using support vector machine. Afr J Comput ICTs. 2015;8:75–82.
18. Qin H, Li X, Yang Z, Shang M. When underwater imagery analysis meets deep learning: a solution at the age of big visual data. In: OCEANS 2015 - MTS/IEEE Washington. 2015; pages 1–5.
19. Rodrigues MTA, Freitas MHG, Pádua FLC, Gomes RM, Carrano EG, Eduardo G. Evaluating cluster detection algorithms and feature extraction techniques in automatic classification of fish species. Pattern Anal Appl. 2015;18(4):783–97.
20. Salman A, Jalal A, Shafait F, Mian A, Shortis M, Seager J, Harvey E. Fish species classification in unconstrained underwater environments based on deep learning. Limnol Oceanogr. 2016;14(9):570–85.
21. Liang J, Qin H, Li X, Zhang C. Deepfish: accurate underwater live fish recognition with a deep architecture. Neurocomputing. 2016;187:49–58.
22. Rathi D, Jain S, Indu S. Underwater fish species classification using convolutional neural network and deep learning. In: 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR). 2017; pages 1–6.
23. Demertzis K, Iliadis LS, Anezakis VD. Extreme deep learning in biosecurity: the case of machine hearing for marine species identification. J Inf Telecommun. 2018;2(4):492–510.
24. Hussain M, Ather I, Wang Z, Ali Z, Riaz S. Automatic fish species classification using deep convolutional neural networks. Wirel Pers Commun. 2019;116:1043–53 (**08**).
25. Chuang M, Hwang J, Williams K. A feature learning and object recognition framework for underwater fish images. IEEE Trans Image Process. 2016;25(4):1862–72.
26. Qiu C, Zhang S, Wang C, Yu Z, Zheng H, Zheng B. Improving transfer learning and squeeze- and-excitation networks for small-scale fine-grained fish image classification. IEEE Access. 2018;6:78503–12.
27. Putra D, Hridayami P, Gede IK, Wibawa KS. Fish species recognition using vgg16 deep convolutional neural network. J Comput Sci Eng. 2019;13(3):124–30.
28. Rawat W, Wang Z. Deep convolutional neural networks for image classification: a comprehensive review. Neural Comput. 2017;29:1–98 (**06**).
29. Canziani A, Paszke A, Culurciello E. An analysis of deep neural network models for practical applications, 2016. arXiv:1605.07678.
30. Deng J, Dong W, Socher R, Li L, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009; pages 248–255.
31. Koustubh. https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception.
32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016; pages 770–778, 06.
33. Boom BJ, Huang PX, He J, Fisher RB. Supporting ground-truth annotation of image datasets using clustering. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR). 2012; pages 1542–1545.
34. Anantharajah K, Ge ZY, McCool C, Denman S, Fookes C, Corke P, Tjondronegoro D, Sridharan S. IEEE winter conference on applications of computer vision. In: IEEE Winter Conference on Applications of Computer Vision. 2014; pages 309–316.
35. Chang CC, Jen C. Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol. 2011;2(3):1–27.

36. Vedaldi A, Fulkerson B. Vlfeat: an open and portable library of computer vision algorithms. In: Proceedings of the 18th ACM International Conference on Multimedia. 2010; pages 1469–1472.
37. Tamou AB, Benzinou A, Nasreddine K, Ballihi L (2018) Underwater Live Fish Recognition by Deep Learning. In: Mansouri A, El Moataz A, Nouboud F, Mammass D (eds) Image and Signal Processing. ICISP 2018. Lecture Notes in Computer Science, vol 10884. Springer, Cham. https://doi.org/10.1007/978-3-319-94211-7_30
38. Zeiler M, Fergus D. Visualizing and understanding convolutional networks. Comput Vis - ECCV. 2014;2014:818–33.
39. Szegedy C, Liu W, Jia Y, Sermanet P, Reed SE, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. Comput Res Reposit. 2014;abs/1409.4842:1–12.