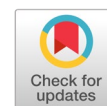


Fish species recognition using transfer learning techniques



Jaisakthi Seetharani Murugaiyan ^{a,1}, Mirunalini Palaniappan ^{b,2,*}, Thenmozhi Durairaj ^{b,3},
Vigneshkumar Muthukumar ^{b,4}

^a Vellore Institute of Technology, Vellore, India

^b Sri Sivasubramaniya Nadar College of Engineering, Chennai, India

¹ jaisakthi.murugaiyan@vit.ac.in; ² miruna@ssn.edu.in; ³ theni_d@ssn.edu.in; ⁴ vigneshkumar14122@cse.ssn.edu.in

* corresponding author

ARTICLE INFO

Article history

Received October 10, 2020

Revised December 5, 2020

Accepted January 6, 2021

Available online July 31, 2021

Keywords

Fish classification

Transfer Learning

SVM Classifier

Deep Neural Network

ABSTRACT

Marine species recognition is the process of identifying various species that help in population estimation and identifying the endangered types for taking further remedies and actions. The superior performance of deep learning for classification is due to the property of estimating millions of parameters that have to be extracted from many annotated datasets. However, many types of fish species are becoming extinct, which may reduce the number of samples. The unavailability of a large dataset is a significant hurdle for applying a deep neural network that can be overcome using transfer learning techniques. To overcome this problem, we propose a transfer learning technique using a pre-trained model that uses underwater fish images as input and applies a transfer learning technique to detect the fish species using a pre-trained Google Inception-v3 model. We have evaluated our proposed method on the Fish4knowledge(F4K) dataset and obtained an accuracy of 95.37%. The research would be helpful to identify fish existence and quantity for marine biologists to understand the underwater environment to encourage its preservation and study the behavior and interactions of marine animals.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. Introduction

Underwater object recognition is one of the most active research fields. Fish population estimation and classification of fish species are essential tasks for ocean observation, assessment of fish stocks and ecosystems, abundance, and diversity in sea /ocean [1]. These tasks benefit scientific and commercial applications like fish farming [2] and understanding food availability and predator-prey relationships [3]. Fish classification is a challenging task due to low-quality underwater videos that are recorded from unconstrained environments (luminosity and illumination changes) in the sea, fish movements, nonlateral fish views, partially visible fish, and the presence of sediments and organic debris [4], [5]. The appearance of fishes in different scales, orientations, curved body shapes, and strong visual similarity between species in shape, size, and coloring further complicates the classification process. Recognition of species from underwater video for assessing fish abundance has limited attention among the researcher's community.

Research on fish recognition has been done using image processing, hand-crafted feature-based [6], machine learning [7], and deep learning [8] techniques. Fish recognition in a constrained environment was proposed in [9] uses the contour matching technique. Larsen *et al.* [10] extracted shape and texture features and classified the fishes using linear discriminant analysis. A deformable template matching method has been proposed using texture and shape for classifying fishes in underwater videos [2], [11]. These methods produce good results only for fishes that were highly distinguishable based on the

patterns. An efficient method for fish detection, counting, and species classification has been done using blob counting and shape analysis [1]. The automatic classification system that extracts the images' geometry, morphology, and texture using a neural network was devised in [12] to classify different species like fish, butterflies, and plants. A sparse representation-based classification (SRC) was proposed by Shiau *et al.* [13] for fish recognition and verification using a maximum probability of partial ranking method. A robust computational method based on an image-set matching approach was proposed using Graph-Embedding Discriminant Analysis [14].

A hierarchical classification level was proposed in [15] that uses the Balance Guaranteed Optimized Tree (BGOT) algorithm for fish classification. Rich feature descriptors and separate features from different fish parts were extracted to improve the discriminative power. A new method of image set classification was presented by Shafait *et al.* [16]. An automatic fish recognition and classification system were proposed by Rodrigues *et al.* [17] using SIFT and PCA for parameterizing the features. Two immunological algorithms, namely Artificial Immune Network and Adaptive Radius Immune Algorithm used these features to cluster fish types. The underwater fish recognition framework was proposed using an unsupervised feature learning technique and an error-resilient classifier [18]. A deep architecture was proposed by Qin *et al.* [19] for fish classification. A special cross-layer pooling approach that combines the features from different layers of the deep network of CNN model was proposed by Shoaib *et al.* [20]. Sparse representation-based classification combined with maximum probability was proposed in [21] based on the visual features of Eigenfaces and Fisherfaces. Alsmadi *et al.* [22] identified the pattern of fishes using a multi-layer feed-forward neural network model with a backpropagation classifier based on color signatures. Vilon *et al.* [5] proposed two supervised methods for fish classification. In the first method, HOG features were extracted, SVM was used for classification, and Google Net architecture was used in the second method. In the 2015 SeaClef contest [15], a task on fish classification was conducted and the best results are achieved through deep learning-based techniques. Siddiqi *et al.* [20] used cross-layer pooling algorithm using pre-trained convolutional neural networks. Salman *et al.* [2] demonstrated the use of CNN for fish species identification.

Fish classification systems based on image processing techniques have a low accuracy rate since the systems do not handle environmental changes, different characteristics, and feature variability. The methods that use machine learning techniques have to undergo segmentation and feature engineering. Due to the unconstrained environment, segmentation may degrade the system performance [23]. Thus, deep learning methods are considered a good choice among all other techniques since they learn automatically from the input images and the system's accuracy depends on the learning parameters. The deep learning methods focus on learning higher-level features automatically by combining several low-level features. These high-level representation gives more information and helps to achieve more performance in classifying the fish types. The automatic feature learning capability resolves the dependency of human-crafted features but requires many training samples. Yet, this approach requires extensive training data.

In order to overcome the problem of data insufficiency, several research works have been reported using transfer learning techniques [24][25]. The transfer learning technique is based on the hypothesis where training images are independent and identically distributed (i.i.d) with the target images. We have proposed a transfer learning technique for fish classification using a pre-trained Google inception v3 model trained using the general images. This model dramatically reduces the requirement of an extensive training dataset and helps in the reduction of the overfitting of a small dataset. Thus, in this paper, we propose and analyze the performance of fish classification using a transfer learning technique that contains three variants. To study the performance of the proposed system, we have used the images obtained from underwater videos captured in an unconstrained environment.

2. Method

We propose a transfer learning technique for classifying the fish species using the images obtained from underwater videos. A pre-trained Google Inception v3 [4] model has been used to extract the weights, and these extracted weights are used to train the input images and classify the test images. In this paper, we have proposed three different methods to classify the fish species. In the first method, we used transfer learning as a feature extractor where representation vectors of the pre-trained model were

used, and the softmax classifier was modified according to the target image classes. In the second method, we have used FineTuning, which is also a transfer learning technique where the first two layers were kept freezing, and the other layers have been trained with target images and in the last method, the weights obtained using the second method were used to train the SVM classifier in order to classify the target images. The performances of all three methods were analyzed.

Transfer learning is an active research area in which the knowledge gained in one or more tasks is applied to different tasks [26]. This can be achieved using a pre-trained model which has been previously trained on a large dataset. The pre-trained model can be customized for a given task which is known as feature extraction and Fine-Tuning. The trained model can be used directly by modifying the output layer, and the rest of the layers can be used to extract features, called as feature extraction model, or the pre-trained model can be trained partially by freezing and unfreezing the required layers based on the application area, known as fine-tuning.

Several pre-trained models such as Google Inception, AlexNet, and VGGNet can be used for image classification based on transfer learning techniques. The Google Inception model contains fully connected layers with global average pooling that averages the channel values across 2D feature maps. Due to this, it achieves more accuracy on the ImageNet dataset with reduced parameters than AlexNet and much faster than VGG. So, a transfer learning-based fish classification system using Google inception model has been proposed in this paper.

2.1. Google Inception V3 as Feature Extractor

The Google Inception v3 model [27] contains 42 layers with inception modules, fully connected layers, and batch normalization layers. Each inception layer consists of convolution layers with different filters followed by the average pooling or max-pooling layers. The convolutional layers consist of learnable filters that learn a hierarchical representation of the input in which the initial layers detect the basic patterns like edges and gradients, while top layers learn patterns specific to the input images. The knowledge gained by the lower layers from the source domain may serve as valuable knowledge for the new domain.

In this method, the Google inception v3 model act as a feature extractor. The representation vectors of the pre-trained model were obtained, and 23 nodes have replaced the softmax layer of v3 to classify 23 different species of the target images. The softmax layer was trained with the target image and the representation vectors obtained from the pre-model, and a model is built. The model is used to predict the classes of the target images. The architecture of the proposed system is depicted in Fig.1.

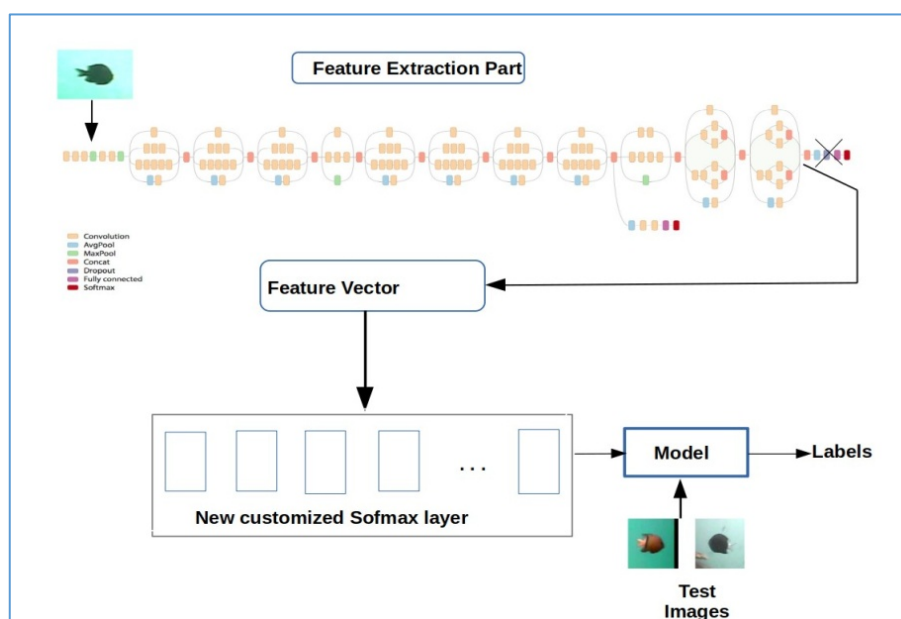


Fig. 1. The architecture of the feature extractor.

2.2. Google Inception v3 as Fine Tuning

In Fine-tuning, the top two inception layers of v3 were kept freezing, and the remaining blocks of the model have been trained with fish images obtained from underwater videos. The SoftMax layer, the last layer of Google inception, has been replaced with 23 nodes and trained with target images. During training, we have used categorical loss and RMS prop optimizer with a learning rate of 0.0001.

2.3. Google Inception v3 with SVM

The representation vectors were obtained by fine-tuning all the layers with target images by excluding the top two layers. The extracted weights were used to train SVM, where the trained model has been used to classify the fish images during testing. We have used non-linear SVM to handle many features and maximize the margin between different data samples. One-Vs-all classification strategy is used to train the SVM. We have used the LibSVM tool [28] to train and built a multiclass SVM classifier. The architecture of the proposed system is depicted in Fig.2.

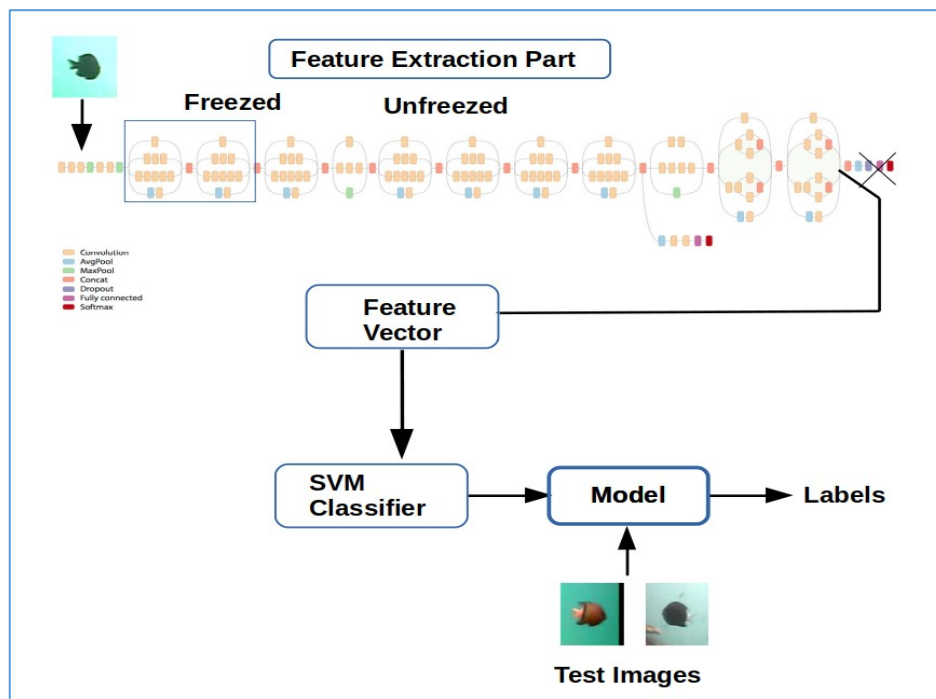


Fig. 2. The architecture of the Google Inception v3 with SVM.

3. Results and Discussion

Fish4knowledge(F4K) dataset [22] is used as the benchmark dataset for fish classification. The dataset consists of 27370 fish images that are obtained from 5824 video clips. These videos are captured using 9 cameras from three different locations under an unconstrained environment. The dataset is annotated with 23 fish types which are more complex and imbalanced. Out of the total images, training and test images are 16430 and 10940, respectively. The intra-class images are different in shape, size, and number of fins. The dataset contains more samples of the most frequent fish class than the infrequent class. The sample images from the dataset are depicted in Fig. 3.

The performance of different transfer learning approaches proposed in this paper were analyzed in terms of accuracy. Accuracy is the ratio of correctly classified fishes and the total number of fishes. We have obtained 79.08% accuracy for the feature extraction method, and for fine-tuning, 90.53% was obtained for the test images. Among these two methods, we have improved accuracy for fine-tuning compared to the feature extraction technique since we have fine-tuned the pre-trained model's top layers. Since the proposed system has learned the basic abstract features of the images from the lower layers, high-level features on the target input images are learned from top layers, leading to high accuracy.

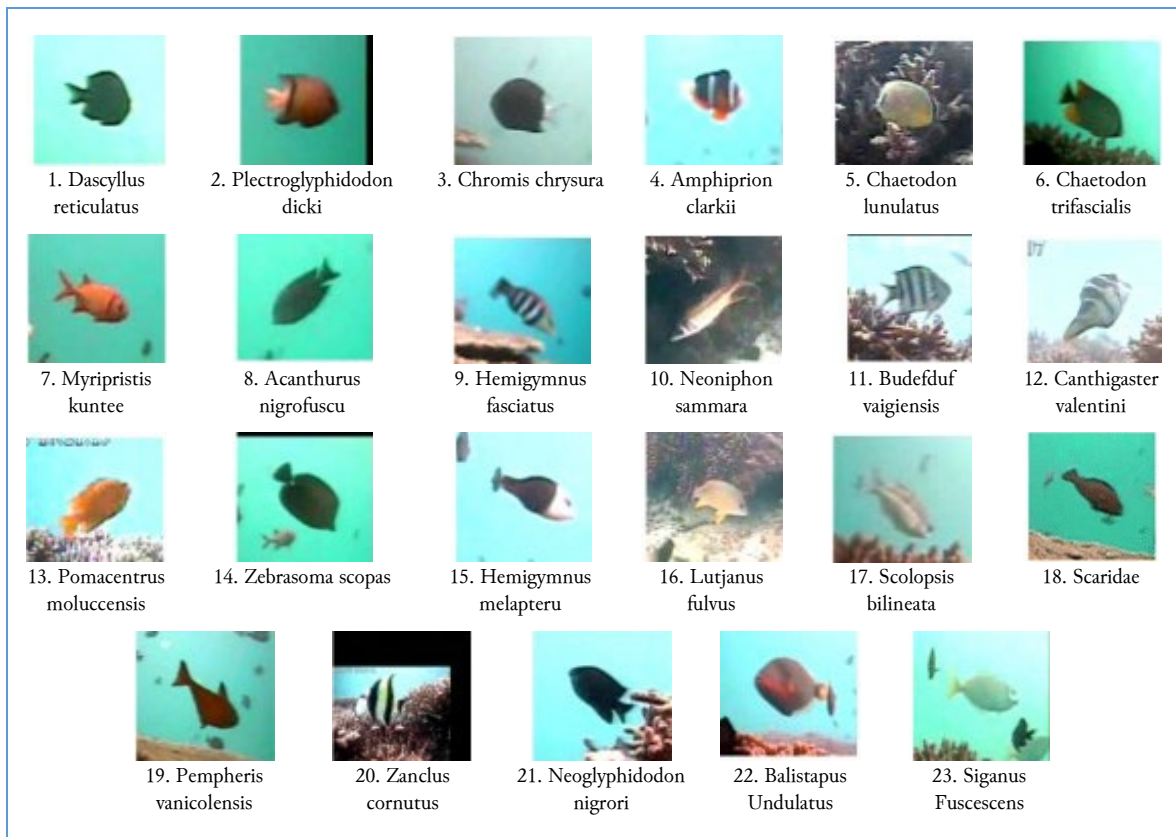


Fig. 3. Sample Images from Fish4knowledge (F4K).

The representation vectors from the training images were obtained through the Google Inception model in the last method. The input images in the dataset are divided into training and test set in the ratio of 60:40. In order to effectively classify the input images, we have fine-tuned the hyperparameters of SVM, namely C and γ , using grid search. The grid search has been performed at two levels. During the first level, a coarse-grained search was made to find the optimal values, and in the next level, the fine-grained search was made locally. We have obtained the C and γ values as 16.0 and 0.03125, respectively. The result of the grid search, along with the accuracies, is depicted in Fig.4.

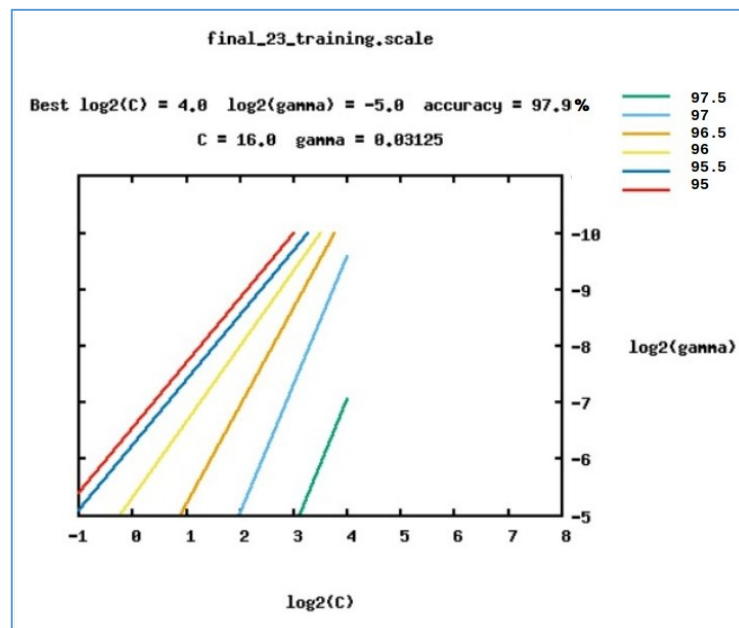


Fig. 4. Search by Libsvm.

We also performed 10-fold cross-validation to validate our training set. We got a training accuracy of 97.9%. The fine-tuned parameters, a multiclass model, were built, and the test images were tested using the model. The input images are classified into 23 different classes. With this method, we have achieved an accuracy of 95.37% for the test set. Classification results obtained for the fish4knowledge dataset for individual species is shown in Table 1. This table shows that the ensemble approach produced improved accuracy because of two levels of training, one using the pre-trained model and the other one using SVM.

Table 1. Accuracy obtained of the proposed Google Inception v3 Model with SVM.

ID	Species	Total No of Images	Correctly Classified	Accuracy(%)
1	Dascyllus Reticulatus	4844	4714	97.32
2	Plectroglyphidodon Dicki	1073	1034	96.36
3	Chromis Chrysur	1437	1358	94.50
4	Amphiprion Clarkii	1619	1609	99.38
5	Chaetodon Lunulatus	1014	1009	99.50
6	Chaetodon Trifascialis	76	59	76.63
7	Myripristis Kuntze	180	152	84.44
8	Acanthurus Nigrofuscu	87	17	19.54
9	Hemigymnus Fasciatus	96	90	93.75
10	Neoniphon Sammara	119	118	99.15
11	Budefdud Vaigiensis	39	22	56.41
12	Canthigaster Valentini	58	31	53.44
13	Pomacentrus Moluccensis	72	47	65.27
14	Zebrasoma Scopas	36	8	22.22
15	Hemigymnus Melapteru	16	8	56.25
16	Lutjanus Fulvus	82	81	98.78
17	Scolopsis Bilineata	19	6	31.58
18	Scaridae	22	8	36.36
19	Pempheris Vanicolensis	11	11	100.00
20	Zanclus Cornutus	8	2	25.00
21	Neoglyphidodon Nigrori	6	1	17.00
22	Balistapus Undulatus	16	3	18.75
23	Siganus Fuscescens	10	3	30.00
Total		10937	10391	95.37

Table 2 depicts the confusion matrix obtained from the multiclass SVM classifier. From the confusion matrix, some species like Acanthurus Nigrofuscu, Zanclus Cornutus, Zebrasoma Scopas, Neoglyphidodon Nigrori, Balistapus Undulatus have high misclassifications because of shallow discriminating power and less number of samples.

We have also evaluated precision, recall, and F1-score for Google Inception v3 Model with SVM which is listed in Table 3. Precision is the measure that projects classifier exactness, whereas recall indicates classifier completeness. F1-score measures the balance between precision and recall. With the proposed method, we have achieved a 70.88% of F1-score, which is a commendable result. Also, the proposed method produced 100% precision for most of the classes, which is better than the existing systems. Among the proposed three methods SVM classifier has achieved high accuracy of 95.37%, and the results are shown in Table 4. SVM classifier produced high accuracy compared to the other methods because SVM is capable of selecting appropriate discriminating features for classification from a large number of features. Moreover, in this approach, parameters are learned both in the pre-trained model and SVM training phases.

Table 2. Confusion Matrix obtained of the Google Inception v3 model with SVM.

class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	4714	35	76	6	4	5	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	30	1034	2	2	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	67	11	1358	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	8	0	0	1609	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	2	0	0	3	1009	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	12	1	0	0	6	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	6	21	1	0	0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	57	3	7	2	0	0	3	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	4	2	0	0	0	0	1	0	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	118	1	0	0	0	0	0	0	0	0	0	0	0	0
11	5	0	0	1	0	2	0	0	5	1	22	3	0	0	0	0	0	0	0	0	0	0	0
12	5	0	0	14	0	0	0	0	8	0	0	31	0	0	0	0	0	0	1	0	1	0	0
13	4	6	0	12	0	3	0	0	0	0	0	0	47	0	0	0	0	0	0	0	0	0	0
14	24	3	1	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
15	6	0	0	1	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	81	0	0	0	0	0	0	0
17	11	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0
18	1	0	3	0	0	0	4	6	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0
20	1	0	0	1	0	1	0	0	3	0	0	0	0	0	0	0	0	0	0	2	0	0	0
21	3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
22	8	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
23	6	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3

Table 3. Precision, Recall, F1-Measure of the Google Inception v3 model with SVM.

ID	Species	Precision	F1-score	Recall
1	Dascyllus reticulatus	95.34	96.31	97.32
2	Plectroglyphidodon dicki	92.32	94.29	96.36
3	Chromis chrysur	93.71	94.10	94.50
4	Amphiprion clarkii	97.37	98.36	99.38
5	Chaetodon lunulatus	99.01	99.25	99.50
6	Chaetodon trifascialis	80.82	78.67	76.63
7	Myripristis kuntee	91.56	87.86	84.44
8	Acanthurus nigrofuscus	72.00	30.74	19.54
9	Hemigymnus fasciatus	81.81	87.37	93.75
10	Neoniphon sammara	99.51	99.51	99.15
11	Budefduf vaigiensis	91.66	69.83	56.41
12	Canthigaster valentini	91.17	67.38	53.44
13	Pomacentrus moluccensis	100.00	78.98	65.27
14	Zebrasoma scopas	100.00	36.36	22.22
15	Hemigymnus melapteru	100.00	72.00	56.25
16	Lutjanus fulvus	100.00	99.38	98.78
17	Scolopsis bilineata	100.00	48.00	31.58
18	Scaridae	100.00	53.32	36.36
19	Pempheris vanicolensis	91.66	95.64	100.00
20	Zanclus cornutus	100.00	40.00	25.00
21	Neoglyphidodon nigrori	50.00	25.373	17.00
22	Balistapus Undulatus	100.00	31.57	18.75
23	Siganus Fuscescens	100.00	46.15	30.00
Over all		88.37	70.88	95.37

We have used three approaches where the first approach is based on the feature extraction, approach two works using a fine-tuning technique, and the last approach used a support vector classifier for classification. The results of our approaches were compared with the results reported in [2][20][29][30]. Hasija *et al.* [14] and Chuang *et al.* [18] used traditional image processing techniques, while Siddiqui *et al.* [20] and Salman *et al.* [2] used deep learning techniques. All the approaches reported their performance on fish species classification in unconstrained environments which is drawn from the Fish4Knowledge dataset. We have used all 23 classes for training and testing purposes, While other approaches listed in Table 4 used only the top 15 species, which contains many samples.

Table 4. Accuracy of proposed methods.

Method	Test Images	Correctly Classified	Mis Classified	Accuracy
Google Inception v3 as Feature extractor	10940	86	2288	79.08%
Google Inception v3 as Fine – Tunning	10940	9904	1036	90.53%
Google Inception v3 with SVM	10940	10391	549	95.37%

Our proposed approaches produced promising results for all the 23 species classification. The SVM classifier of our approach produced better results when compared to all other approaches despite fewer samples for few classes. The SVM classifier acts as a good feature selector that helps select prominent features for each species, which helps SVM produce good accuracy. In our second approach, the hierarchical feature learning capability of the CNN layers present in the Google inception model helps to capture the dependent visual features to produce better classification results. Even though our first approach used the Google Inception model, we obtained less accuracy because it used pre-trained weights of the ImageNet data. Table 5. shows the comparison of our proposed methods with the other methods reported in the literature.

Table 5. Comparison with existing works approaches accuracy.

Approach	Accuracy
Hasija <i>et al.</i> [14]	91.66%
Chuang <i>et al.</i> [18]	93.80%
Siddiqui <i>et al.</i> [20] (15 species)	96.73%
Salman <i>et al.</i> [2] (15 species)	87.46%
Google Inception v3 with SVM (23 species)	95.37%
Google Inception v3 as Fine Tuning (23 species)	90.53%

4. Conclusion

We have proposed three different approaches for classifying fish species using the transfer learning technique. The first method acts as a feature extractor, extracts features using the Google Inception-v3 model, and classifies all the test images using the SoftMax classifier. The second method is based on fine-tuning that tunes all the layers of the Google Inception model using input images except the first two layers. We have achieved an accuracy of 90.53% by this method. The third method, ensembling both Google Inception v3 and SVM classifier, attains an accuracy of 95.37%. We have compared the performance of our approaches with the state of art approaches, and it is observed that our transfer learning-based approach using SVM classifier performs better than existing work. As a feature work, we would like to use a deep learning model to classify the fishes.

Acknowledgment

The authors would like to thank the management of VIT University, Vellore, India, and SSN College of Engineering, Chennai India, for funding the respective research labs where the research work is being carried out. One of the authors, S M Jaisakthi, would like to thank NVIDIA for providing a GPU grant in support of this research work, and similarly, P Mirunalini would like to thank the management for providing the GPU machine where this research is carried out.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

- [1] J. N. Fabic, I. E. Turla, J. A. Capacillo, L. T. David, and P. C. Naval, "Fish population estimation and species classification from underwater video sequences using blob counting and shape analysis," in *2013 IEEE International Underwater Technology Symposium (UT)*, 2013, pp. 1–6, doi: [10.1109/UT.2013.6519876](https://doi.org/10.1109/UT.2013.6519876).
- [2] A. Salman *et al.*, "Fish species classification in unconstrained underwater environments based on deep learning," *Limnol. Oceanogr. Methods*, vol. 14, no. 9, pp. 570–585, Sep. 2016, doi: [10.1002/lom3.10113](https://doi.org/10.1002/lom3.10113).
- [3] M. R. Heithaus and L. M. Dill, "Food availability and tiger shark predation risk influence bottlenose dolphin habitat use," *Ecology*, vol. 83, no. 2, pp. 480–491, 2002, doi: [10.1890/0012-9658\(2002\)083\[0480:FAATSP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[0480:FAATSP]2.0.CO;2).
- [4] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9, doi: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594).
- [5] S. Villon, M. Chaumont, G. Subsol, S. Villéger, T. Claverie, and D. Mouillot, "Coral Reef Fish Detection and Recognition in Underwater Videos by Supervised Machine Learning: Comparison Between Deep Learning and HOG+SVM Methods," 2016, pp. 160–171. doi: [10.1007/978-3-319-48680-2_15](https://doi.org/10.1007/978-3-319-48680-2_15)
- [6] X. Wang, J. Ouyang, D. Li, and G. Zhang, "Underwater Object Recognition Based on Deep Encoding-Decoding Network," *J. Ocean Univ. China*, vol. 18, no. 2, pp. 376–382, Apr. 2019, doi: [10.1007/s11802-019-3858-x](https://doi.org/10.1007/s11802-019-3858-x).
- [7] G. Liang, F. Chen, Y. Liang, Y. Feng, C. Wang, and X. Wu, "A Manufacturing-Oriented Intelligent Vision System Based on Deep Neural Network for Object Recognition and 6D Pose Estimation," *Front. Neurobot.*, vol. 14, Jan. 2021, doi: [10.3389/fnbot.2020.616775](https://doi.org/10.3389/fnbot.2020.616775).
- [8] J. Hu, D. Zhao, Y. Zhang, C. Zhou, and W. Chen, "Real-time nondestructive fish behavior detecting in mixed polyculture system using deep-learning and low-cost devices," *Expert Syst. Appl.*, vol. 178, p. 115051, Sep. 2021, doi: [10.1016/j.eswa.2021.115051](https://doi.org/10.1016/j.eswa.2021.115051).
- [9] D.-J. Lee, R. B. Schoenberger, D. Shiozawa, X. Xu, and P. Zhan, "Contour matching for a fish recognition and migration-monitoring system," in *Two-and Three-Dimensional Vision Systems for Inspection, Control, and Metrology II*, 2004, vol. 5606, pp. 37–48. doi: [10.1117/12.571789](https://doi.org/10.1117/12.571789)
- [10] R. Larsen, H. Olafsdottir, and B. K. Ersbøll, "Shape and Texture Based Classification of Fish Species," 2009, pp. 745–749. doi: [10.1007/978-3-642-02230-2_76](https://doi.org/10.1007/978-3-642-02230-2_76)
- [11] N. J. C. Strachan, "Length measurement of fish by computer vision," *Comput. Electron. Agric.*, vol. 8, no. 2, pp. 93–104, Mar. 1993, doi: [10.1016/0168-1699\(93\)90009-P](https://doi.org/10.1016/0168-1699(93)90009-P).
- [12] Y.-H. Hsiao, C.-C. Chen, S.-I. Lin, and F.-P. Lin, "Real-world underwater fish recognition and identification, using sparse representation," *Ecol. Inform.*, vol. 23, pp. 13–21, Sep. 2014, doi: [10.1016/j.ecoinf.2013.10.002](https://doi.org/10.1016/j.ecoinf.2013.10.002).

- [13] Y.-H. Shiau, F.-P. Lin, and C.-C. Chen, "Using sparse representation for fish recognition and verification in real world observation," in *Int. Workshop Visual Observ. Anal. Animal Insect Behav. (VAIB), conjunction Int. Conf. Pattern Recognit*, 2012. Available at: [Google Scholar](#).
- [14] S. Hasija, M. J. Buragohain, and S. Indu, "Fish Species Classification Using Graph Embedding Discriminant Analysis," in *2017 International Conference on Machine Vision and Information Technology (CMVIT)*, 2017, pp. 81–86, doi: [10.1109/CMVIT.2017.23](#).
- [15] A. Joly *et al.*, "LifeCLEF 2015: Multimedia Life Species Identification Challenges," 2015, pp. 462–483. doi: [10.1007/978-3-319-24027-5_46](#)
- [16] F. Shafait *et al.*, "Fish identification from videos captured in uncontrolled underwater environments," *ICES J. Mar. Sci. J. du Cons.*, vol. 73, no. 10, pp. 2737–2746, Nov. 2016, doi: [10.1093/icesjms/fsw106](#).
- [17] M. T. A. Rodrigues, F. L. C. Padua, R. M. Gomes, and G. E. Soares, "Automatic fish species classification based on robust feature extraction techniques and artificial immune systems," in *2010 IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, 2010, pp. 1518–1525, doi: [10.1109/BICTA.2010.5645273](#).
- [18] M.-C. Chuang, J.-N. Hwang, and K. Williams, "A Feature Learning and Object Recognition Framework for Underwater Fish Images," *IEEE Trans. Image Process.*, pp. 1–1, 2016, doi: [10.1109/TIP.2016.2535342](#).
- [19] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, "DeepFish: Accurate underwater live fish recognition with a deep architecture," *Neurocomputing*, vol. 187, pp. 49–58, Apr. 2016, doi: [10.1016/j.neucom.2015.10.122](#).
- [20] S. A. Siddiqui *et al.*, "Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data," *ICES J. Mar. Sci.*, vol. 75, no. 1, pp. 374–389, Jan. 2018, doi: [10.1093/icesjms/fsx109](#).
- [21] P. X. Huang, B. J. Boom, and R. B. Fisher, "Underwater Live Fish Recognition Using a Balance-Guaranteed Optimized Tree," 2013, pp. 422–433. doi: [10.1007/978-3-642-37331-2_32](#)
- [22] M. K. Alsmadi, K. B. Omar, and S. A. M. Noah, "Fish classification based on robust features extraction from color signature using backpropagation classifier," *J. Comput. Sci.*, vol. 7, no. 1, p. 52, 2011. doi: [10.3844/jcssp.2011.52.58](#)
- [23] R. Raj, N. D. Londhe, and R. Sonawane, "Automated psoriasis lesion segmentation from unconstrained environment using residual U-Net with transfer learning," *Comput. Methods Programs Biomed.*, vol. 206, p. 106123, Jul. 2021, doi: [10.1016/j.cmpb.2021.106123](#).
- [24] Y. Ren, X. Chen, S. Wan, K. Xie, and K. Bian, "Passenger Flow Prediction in Traffic System Based on Deep Neural Networks and Transfer Learning Method," in *2019 4th International Conference on Intelligent Transportation Engineering (ICITE)*, 2019, pp. 115–120, doi: [10.1109/ICITE.2019.8880220](#).
- [25] J. Duan, Y. He, and X. Wu, "Serial transfer learning (STL) theory for processing data insufficiency: Fault diagnosis of transformer windings," *Int. J. Electr. Power Energy Syst.*, vol. 130, p. 106965, Sep. 2021, doi: [10.1016/j.ijepes.2021.106965](#).
- [26] N. Hernandez, J. Lundström, J. Favela, I. McChesney, and B. Arnrich, "Literature Review on Transfer Learning for Human Activity Recognition Using Mobile and Wearable Devices with Environmental Technology," *SN Comput. Sci.*, vol. 1, no. 2, p. 66, Mar. 2020, doi: [10.1007/s42979-020-0070-4](#).
- [27] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826, doi: [10.1109/CVPR.2016.308](#).
- [28] B. J. Boom, P. X. Huang, J. He, and R. B. Fisher, "Supporting ground-truth annotation of image datasets using clustering," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 2012, pp. 1542–1545. Available at: [Google Scholar](#).
- [29] L. E. Greene and W. S. Alevizon, "Comparative accuracies of visual assessment methods for coral reef fishes," *Bull. Mar. Sci.*, vol. 44, no. 2, pp. 899–912, 1989. Available at: [Google Scholar](#).
- [30] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011, doi: [10.1145/1961189.1961199](#).