A Robust Fish Species Classification Framework: FRCNN-VGG16-SPPNet

Mei-Hsin Chen
Feng Chia University
Ting-Hsuan Lai
Feng Chia University
Yao-Chung Chen (brucechen@gis.tw)
Feng Chia University
Tien-Yin Chou
Feng Chia University
Fang-Shii Ning
National Chengchi University

Research Article

Keywords: Classification, FRCNN, VGG16, SPPNet

Posted Date: April 19th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2825927/v1

Abstract

This study proposes a novel framework for fish species classification that combines FRCNN (Faster Region-based Convolutional Neural Network), VGG16 (Visual Geometry Group 16), and SPPNet (Spatial Pyramid Pooling network). The proposed FRCNN-VGG16-SPPNet framework combines the strengths of FRCNN's fast object detection and localization, VGG16's convenient transfer learning and fast classification performance, and SPPNet's image processing flexibility and robustness in handling input images of any size. First, FRCNN is used to detect and extract target objects from images containing multiple objects. Subsequently, photos of various fish species at different scales are fed into VGG16-SPPNet, which performs basic feature extraction using transfer learning theory. SPPNet further processes the input images by performing pooling operations of different scales. Finally, VGG16 identifies important features to perform object classification. The proposed framework achieves higher accuracy compared to traditional single VGG16 models, particularly in classifying objects of different sizes, with an accuracy rate of 0.9318, which is 26% higher than traditional single VGG16 models. The proposed framework is efficient, convenient, reliable, and robust for object classification and has potential for various applications in image recognition and classification.

1 Introduction

Fish are an essential food source for millions of people worldwide, particularly for coastal communities. However, overfishing, illegal fishing, and other human activities are threatening the sustainability of marine fish stocks[1, 2]. As a result, the United Nations has included SDG 14 among the Sustainable Development Goals (SDGs) to conserve and sustainably use marine resources[3, 4]. Achieving SDG 14 is critical to the health of ocean ecosystems, the livelihoods of those who depend on them, and the global economy[5]. Sustainable fisheries management is crucial to prevent overfishing and promote the long-term sustainability of marine fish stocks. This includes setting appropriate catch limits, reducing bycatch, implementing effective monitoring and enforcement measures, and establishing marine protected areas (MPAs)[6, 7]. Sustainable fisheries management can also enhance the economic benefits associated with fisheries by ensuring their long-term sustainability. According to the Food and Agriculture Organization (FAO), the total value of global fish production was estimated at USD 401 billion in 2016[8], and the fisheries sector employed more than 50 million people[9]. Implementing sustainable fish-eries management practices can ensure the long-term sustainability of the fisheries sector and support the livelihoods of those who depend on it. Preventing overfishing and im-plementing sustainable fisheries management practices are important means of achieving SDG 14 and promoting the long-term sustainability of marine fish stocks[10]. Govern-ments, fisheries managers, and other stakeholders must work together to implement these measures to ensure a sustainable future for our oceans and the people who depend on them. Taiwan is an island country surrounded by the sea on all sides, with a tropical and subtropical climate. It has rich marine ecology and is also an important habitat for many marine organisms, creating considerable biological resources. As the public gradually pays more attention to the concept of environmental protection and ecological symbiosis, marine conservation has become a public affair that requires the participation of all citi-zens to maintain it. Therefore, the purpose of this research is to establish a novel, fast, convenient, and accurate fish identification framework. This will allow citizens to identify and record fish species through the framework while fishing for leisure, thereby promot-ing the conservation of marine biodiversity.

In recent years, the rapid development of computer vision, deep learning, and infor-mation technology has led to the use of AI technology in fish species classification[11, 12]. Researchers are committed to developing novel classification models, which can be cate-gorized into three types: machine learning-based, deep learning-based, or hybrid methods.

Machine learning-based methods typically employ supervised and unsupervised learning[13]. In fish species classification, supervised learning methods use labeled image datasets to train a model to identify fish species. This approach requires human selection of fish features, human experience, and knowledge, and has a high requirement for data quality. While this method has high interpretability and makes it easier to understand how the model makes predictions, it also requires manual feature extraction, which is la-bor-intensive, and the selection of features can affect the performance of the model. Com-monly used algorithms include Support Vector Machine (SVM)[14–17], Artificial Neural Network (ANN)[18, 19], Decision Tree (DT)[17, 20], Random Forest (RF)[16, 17, 21], K-Nearest Neighbors (KNN)[16, 17, 22], and Logistic Regression[16, 23].

Deep learning methods are a type of machine learning method based on neural net-works, which automatically learn features for classification. Compared with machine learning, deep learning requires more data and more powerful computing capabilities. However, it has the advantage of automatic feature extraction, reducing manual work and enabling the handling of more complex features. Its disadvantage is higher black-box na-ture, making it difficult to understand the internal workings of the model, and it requires more computing resources. Commonly used models include VGG16[24, 25], AlexNet[26, 27], GoogleNet[12, 28], ResNet50[29–31], MobileNet[29, 32], and Mask Re-gion-based Convolutional Neural Network (Mask R-CNN)[33–35]. These methods are all classification models derived from CNN. By using transfer learning, the training of the model can be significantly accelerated, and its accuracy can be improved.

Hybrid classification methods refer to methods that combine machine learning and deep learning techniques, fully utilizing their respective advantages and avoiding their disadvantages to improve the accuracy and efficiency of fish species classification. Ma-chine learning methods are usually suitable for small samples and low-dimensional data, while deep learning methods are suitable for large samples and high-dimensional data. Therefore, this hybrid method can handle various fish species classification problems. Commonly used algorithms include CNN-SVM[36–38], CNN-KNN[30, 37, 39], CNN-ANN[40].

As it is obvious, deep learning and transfer learning techniques are the current re-search focus in fish species classification. Fine-tuning pre-trained models has been found to be an effective approach in adapting to new classification tasks. VGG16, a classic con-volutional neural network model, achieved outstanding performance in the ImageNet image classification competition and has been widely applied in various image classifica-tion tasks[41–46], Therefore, VGG16 was selected as the backbone framework in this study. However, irrelevant features in images can affect classification accuracy. Ideally, limiting images to only contain relevant target object information can improve model performance. Furthermore, pre-trained models are fixed in terms of input image size, two issues arise: the potential loss of important information due to scaling, and the significant computa-tion and storage resources required to handle a large number of images of varying sizes. which limits their application range. As the proposed classification framework in this study aims to provide public access, image sources are limited and their sizes are incon-sistent. To address these issues, the Faster Region-based Convolutional Neural Network (FRCNN) was introduced for fish detection and localization in raw images, and the Spa-tial Pyramid Pooling network (SPPNet), which incorporates a Spatial Pyramid Pooling (SPP) layer, was utilized to process input images of arbitrary sizes and generate fixed-size feature representations. This approach can avoid image distortion or information loss caused by image resizing. Therefore, a novel, convenient, and highly accurate hybrid clas-sification model, FRCNN-VGG16-SPPNet, was proposed in this study.

2 Dataset And Methods

2.1 Imagery Collection and Preprocess

In this study, we utilized image data from five prevalent marine fish species in Tai-wan as the source for training and testing a classification model. A total of 345 images was collected, including Pomadasys argenteus, Mugil cephalus, Acanthopagrus latus, Carangoides hedlandensis, and Caranx sexfasciatus (as shown in Fig. 1). Each species had approximately 54–75 images collected. The images were sourced from databases such as FishBase[47] and ImageNet[48].

To effectively address the issue of imbalanced data and enhance the ability to recog-nize minority categories, as well as increase the diversity of training samples, improve the generalization and stability of the classifier, and prevent overfitting[49], we adopted the Image Augmentation technique[50] to perform operations such as flipping, rotation, scal-ing, translation, cropping, adjusting contrast, brightness, and adding noise to the original images to generate similar but not identical images (as shown in Fig. 2). After the image augmentation, a total of 1,690 images was obtained, the number of images for each fish species ranged from 270 to 370 (as shown in Table 1). In this study, 80% of the image da-tasets for each fish species was used for model training, while the remaining 20% was used for model validation.

Table 1
Sample sizes used for the proposed classification model

Species	Number of Collected Images	Total Number of Images after Augmentation
Pomadasys argenteus	54	270
Mugil cephalus	71	350
Acanthopagrus latus	74	355
Carangoides hedlandensis	71	345
Caranx sexfasciatus	75	370
Total	345	1,690

2.2 Proposed FRCNN-VGG16-SPPNet Framwork

CNN is a type of deep learning-based neural network[51] that is widely used in image recognition, computer vision, and natural language processing[52]. The main feature of CNN is that they can automatically learn and extract relevant features from raw data, such as images, without the need for manual feature engineering. The architecture of CNN consists of convolutional layers, pooling layers, and fully connected layers. CNN has achieved state-of-the-art performance in various computer vision tasks, such as image classification, object detection, and segmentation. In this study, the three highly performing image detection, classification, and recognition techniques integrated - FRCNN, VGG16, and SPPNet - are all derived from CNN[52–54].

FRCNN, a state-of-the-art CNN-based object detection model that uses a Region Proposal Network (RPN) to generate region proposals and a subsequent FRCNN network to perform object classification and bounding box regression[55]. FRCNN has several advantages over previous object detection methods such as its ability to accurately detect and classify objects in images with high variability and clutter, and its relatively fast training and inference time. FRCNN has been applied to a wide range of applications, including pedestrian detection, face detection, vehicle detection, and object recognition in natural scenes. In addition, FRCNN has been shown to be effective in detecting and classifying objects in various field imagery, such as satellite, aerial, medical, biological imagery. One of the key benefits of FRCNN is its ability to handle a large number of object classes, making it suitable for applications with many different types of objects. Furthermore, FRCNN can be used in real-time applications, such as video surveillance and autonomous driving, due to its fast processing speed.

VGG16, a deep CNN model that has achieved leading-edge results in various image classification tasks. It has several advantages over other image classification models, including its deep architecture, which allows it to learn highly complex and abstract features from input images. VGG16 has also been shown to have good generalization performance, meaning that it can accurately classify images from previously unseen classes[56]. VGG16 has been applied to a wide range of image classification tasks. The success of VGG16 in fish species classification has important applications in the field of marine biology and fisheries management, where accurate identification and monitoring of fish populations is critical for conservation and sustainable fisheries practices. Its ability to learn highly complex features from images has made it a popular choice for these types of applications. Another advantage of VGG16 is its availability in popular deep learning libraries such as Keras, TensorFlow, and matlab, making it easy to implement and train for a wide range of image classification tasks. Furthermore, the pre-trained weights of the VGG16 model can be used for transfer learning, where the model can be fine-tuned on a smaller dataset for a specific classification task.

SPPNet, or Spatial Pyramid Pooling Network, was proposed to address the problem of varying input sizes in image classification tasks. SPPNet has several advantages over traditional CNN model, including its ability to handle input images of varying sizes without the need for cropping or resizing[57]. This is achieved by using spatial pyramid pooling layers that allow the model to extract features at multiple scales and resolutions, making it more robust to changes in input image sizes. SPPNet has been applied to various image classification tasks, including object recognition, facial recognition, and scene understanding. In particular, SPPNet has been used in image classification tasks where input images have varying sizes or where a large number of features need to be extracted from the input images. Its ability to handle varying input sizes and extract features at multiple scales has made it a popular choice for these types of applications. Another advantage of SPPNet is its ability to reduce the number of parameters in the model without sacrificing accuracy. This is achieved by using spatial pyramid pooling layers to extract features, which allows the model to learn from a smaller number of parameters while still achieving high accuracy. This makes SPPNet a more efficient and scalable model compared to traditional CNN model.

Without a doubt, each of the aforementioned image recognition algorithms has its own advantages and applicable conditions. The purpose of this study is to propose a fast, flexible, and accuracy fish species classification framework that is not limited by the image size of transfer learning node. Therefore, a hybrid model, FRCNN-VGG16-SPPNet, was constructed by utilizing the advantages of the three CNN-based algorithms. Firstly, the outstanding object detection and positioning characteristics of FRCNN were used to search for fish in photos and crop the fish images with the most suitable rectangular size based on their shape and size. These cropped images were then inputted into the VGG16 image classifier, which removed non-essential object information, presenting the fish as the main subject to improve classification speed and accuracy. As VGG16 is a pre-trained model, good classification accuracy can still be achieved even with a small number of sample images, using the theory of transfer learning. Therefore, VGG16 is fed with the images extracted by FRCNN for the basic feature extraction of fish species. Finally, because the input image size of VGG16 was set to 224x244 RGB images, while the size of the images captured in the FRCNN stage was inconsistent, SPPNet was set before the fully connected layer. The input image was divided into multiple scales by the spatial pyramid pooling layer, and each scale image was pooled to obtain a fixed-size feature vector. In this way, different sized images can obtain the same size of feature vectors, reducing the CNN network's sensitivity to image size and improving the network's generalization performance. Figure 3 and Table 2 present the workflow and hyperparameters of the FRCNN-VGG16-SPPNet framework.

Table 2
Hyperparameters of proposed FRCNN-VGG16-SPPNet model

Parameters	**Value**
Kernel size of CNN	3×3
Batch size	32
Epochs	200
Optimizer	sgd
momentum	0.5
decay	1e-7
nesterov	False

2.3 Experimental Equipment and Assessment Indicators

The experiments were processed on a desktop computer equipped with Intel Core i7-9700F processor with 32 GB RAM, Nvidia RTX 2070 SUPER graphical card, Ubuntu 20.04. The model verification process was divided into training and independent testing. The proposed FRCNN-VGG16-SPPNet method's classification performance was compared with the single VGG16 model employing classic evaluation indicators, such as Accuracy, Precision, Recall, and F1Score. The relevant assessment indicators are expressed as follows[58] (Fig. 4, Equations (1) to (4)):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F}_1\text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

3 Dataset And Methods

To evaluate the effectiveness of the proposed FRCNN-VGG16-SPPNet model, this study compares it with the widely used single VGG16 model and assesses its performance using metrics such as Confusion Matrix, Precision, Recall, F1-Score, Accuracy, and Learning Curves. Figure 5 illustrates the Confusion

Matrix for both models, while Table 3 presents the performance of VGG16 and FRCNN-VGG16-SPPNet for each species during the training stage. The FRCNN-VGG16-SPPNet model exhibits superior training results, with Precision, Recall, F1-Score, and Accuracy all reaching 0.9993, compared to the single VGG16 model, which obtains Precision, Recall, F1-Score, and Accuracy of 0.9754, 0.9748, 0.9747, and 0.9749, respectively. The findings demonstrate that both models demonstrate excellent data fitting ability, however, FRCNN-VGG16-SPPNet's overall performance exceeds that of the single VGG16 model by approximately 2.06%.

Table 3
Performance evaluation indicators for model training

Method	Species	Precision	Recall	F1-Score	Accuracy
VGG16	Pomadasys argenteus	0.9811	0.9630	0.9720	-
	Mugil cephalus	0.9824	0.9964	0.9894	-
	Acanthopagrus latus	0.9790	0.9859	0.9825	-
	Carangoides hedlandensis	0.9416	0.9928	0.9665	-
	Caranx sexfasciatus	0.9928	0.9358	0.9635	-
	All	0.9754	0.9748	0.9747	0.9749
The proposed FRCNN-VGG16-SPPNet	Pomadasys argenteus	1.0000	1.0000	1.0000	-
	Mugil cephalus	1.0000	1.0000	1.0000	-
	Acanthopagrus latus	1.0000	1.0000	1.0000	-
	Carangoides hedlandensis	1.0000	0.9964	0.9982	-
	Caranx sexfasciatus	0.9966	1.0000	0.9983	-
	All	0.9993	0.9993	0.9993	0.9993

A robust classification model is characterized by its ability to produce accurate and reliable predictions that generalize well to new data. In addition to classification accuracy, generalizability is a critical attribute of a reliable model, whereby it is capable of performing well on data that it has not encountered during the training phase. To evaluate a model's reliability, consistency in performance during training and validation is an important indicator. Typically, a model's training accuracy will exceed its validation accuracy, but a large discrepancy may indicate overfitting, which could result in overestimation and erroneous predictions. To assess a model's generalizability, independent testing using unseen data is an essential step in machine learning model development and evaluation, which enables unbiased evaluation of the model's performance.

Independent testing is a critical component in model validation, as it serves to confirm the model's generalization capacity, i.e., its ability to perform well on novel and unseen data. Furthermore, independent testing can mitigate data leakage concerns and diminish model selection errors. In the absence of an independent testing dataset, the validity of the validation outcomes may be compromised, leading to overfitting or underfitting. Notably, Table 4 demonstrates that the proposed FRCNN-VGG16-SPPNet exhibited superior testing results, with Precision, Recall, F1-Score, and Accuracy reaching 0.9382, 0.9260, 0.9294, and 0.9318, respectively. Conversely, the single VGG16 model obtained only 0.7430, 0.7350, 0.7323, and 0.7396, respectively. Upon further comparison of the training and testing results of the classification models, it was observed that the VGG16 model demonstrated a significant variation in Precision, Recall, F1-Score, and Accuracy, ranging from 0.2324 to 0.2424. Conversely, the proposed FRCNN-VGG16-SPPNet model exhibited a notably narrower range of differences, from 0.0611 to 0.0733, indicative of a comparatively consistent performance and superior generalization ability. Notably, these findings show that the VGG16 model may be subject to overfitting, while the proposed FRCNN-VGG16-SPPNet model offers enhanced robustness, reliability and stability.

Table 4
Performance validation indicators for model testing

Method	Species	Precision	Recall	F1-Score	Accuracy
VGG16	Pomadasys argenteus	0.6735	0.6111	0.6408	-
	Mugil cephalus	0.8684	0.9429	0.9041	-
	Acanthopagrus latus	0.6782	0.8310	0.7468	-
	Carangoides hedlandensis	0.6486	0.6957	0.6713	-
	Caranx sexfasciatus	0.8462	0.5946	0.6984	-
	All	0.7430	0.7350	0.7323	0.7396
The proposed FRCNN-VGG16-SPPNet	Pomadasys argenteus	1.0000	0.8113	0.8958	-
	Mugil cephalus	0.9333	1.0000	0.9655	-
	Acanthopagrus latus	0.9577	0.9577	0.9577	-
	Carangoides hedlandensis	0.8553	0.9420	0.8966	-
	Caranx sexfasciatus	0.9444	0.9189	0.9315	-
	All	0.9382	0.9260	0.9294	0.9318

Table 5 presents the performance improvement rates of the proposed hybrid model in comparison to the conventional approach of utilizing only VGG16 for classification. The metrics used to evaluate the performance include Precision, Recall, F1-Score, and Accuracy, which were observed to have improved by 26.27%, 25.99%, 26.92%, and 25.99%, respectively. The results highlight the superiority of the proposed hybrid model over the traditional method. Figure 6 and Fig. 7 provide a graphical representation of the learning curves of the two models. The proposed FRCNN-VGG16-SPPNet model demonstrates consistent training and testing results, indicating its stability and generalizability. Moreover, the model exhibits a fast convergence rate with a commendable classification accuracy. These findings suggest that the proposed FRCNN-VGG16-SPPNet model is capable of effective fish species classification.

In conclusion, the results of the study affirm the superior performance of the proposed hybrid model over the conventional VGG16 approach. The findings provide empirical support for the potential of the FRCNN-VGG16-SPPNet model as a robust tool for accurate fish species classification.

Table 5
Improvement rate for the proposed FRCNN-VGG16-SPPNet model

Method	**Species**	**FRCNN-VGG16-SPPNet Improvement Rate (%)**			
		Precision	**Recall**	**F1-Score**	**Accuracy**
VGG16	Pomadasys argenteus	48.48	32.76	39.79	-
	Mugil cephalus	7.47	6.06	6.79	–
	Acanthopagrus latus	41.21	15.25	28.24	-
	Carangoides hedlandensis	31.87	35.40	33.56	
	Caranx sexfasciatus	11.60	54.54	33.38	-
	All	26.27	25.99	26.92	25.99

Improvement Rate (%) =
(The proposed FRCNN-VGG16-SPPNet model - VGG16 model) / VGG16 model×100

The performance evaluation of a classification model is influenced by a multitude of factors, while the model's design is tailored to the specific requirements of its users. The aim of this study is to develop a model architecture that is convenient, reliable, stable, and highly accurate, for use on mobile devices in fish species recognition by both the general public and marine conservationists. Two major factors that significantly affect the classification accuracy of machine learning models are data quality and feature selection/extraction. Poor data quality, including data skewness, noise, imbalanced samples, and missing values, can negatively impact model training and hinder the learning of effective features and patterns from the data. Thus, it is necessary to perform data preprocessing and cleaning to improve data quality prior to model training. Additionally, feature selection and extraction are critical factors that directly influence the model's classification ability. The ability to select and extract effective features can enhance the model's classification performance, whereas inappropriate feature selection or failure to extract crucial features from the data can result in poor classification outcomes. Appropriate methods for feature selection and extraction must be chosen based on the data's characteristics to enhance the model's classification ability.

In this study, we propose the FRCNN-VGG16-SPPNet framework, which integrates algorithms with unique advantages in image object detection and localization, classification, and feature vector transformation. FRCNN and SPPNet play crucial roles in this framework and provide a multiplying effect that effectively enhances the performance of conventional single VGG16 models. FRCNN automatically detects fish in images containing other objects and crops images centered on the fish, significantly improving the image's recognizability and reducing the complexity of model training and classification. SPPNet's Spatial Pyramid Pooling can transform images of various sizes into feature vectors of the same size, addressing the fixed input image size issue in VGG16 Transfer Learning technology and enabling effective processing of images of different sizes, which is more convenient for model developers and users.

Based on the research analysis presented in this chapter, it is evident that the proposed FRCNN-VGG16-SPPNet framework can highlight the image features of target objects, handle images of different sizes, and exhibit exceptional classification performance.

4 Conclusions

Hybrid methods are currently a prominent approach for image classification, whereby the strengths of different algorithms are combined to enhance classification performance. In this study, the findings provide compelling evidence to support this assertion. Specifically, the study proposes a classification framework that has been successfully deployed in government departments in Taiwan. Users can conveniently upload images to a cloud server via their mobile devices for computation and subsequently obtain the results of fish species classification. Notably, the study focuses on the overall performance of the classification model and does not analyze differences in individual fish species classification accuracy. Nevertheless, the primary factor that potentially causes differences may stem from the quality of the images, given that the images used in the study were obtained from diverse internet sources that can vary in quality.

The framework proposed in this study can serve as a reference for other object classification tasks. Furthermore, future studies will explore the performance of different pre-trained models and incorporate novel object detection and localization methods such as YOLOv8 to develop new approaches continually for the classification model. Overall, the present study contributes to the existing body of knowledge on image classification by highlighting the advantages of hybrid methods and presenting a practical classification framework that can be used in real-world applications.

Declarations

Author Contributions: Conceptualization, M.-H.C. and T.-H.L.; methodology, M.-H.C., T.-H.L. and Y.-C.C.; software, M.-H.C., Y.-C.C. and T.-Y.C.; validation, M.-H.C., Y.-C.C., T.-Y.C. and F.-S.N.; formal analysis, M.-H.C., T.-H.L., Y.-C.C. and T.-Y.C.; investigation, M.-H.C., T.-H.L., Y.-C.C. and T.-Y.C.; resources, M.-H.C. and T.-Y.C.; data curation, M.-H.C., T.-H.L. and T.-Y.C.; writing—M.-H.C., Y.-C.C. and T.-H.L.; writing—review and editing, M.-H.C., Y.-C.C., T.-Y.C. and F.-S.N. All authors have read and agreed to the published version of the manuscript.

Funding: Not applicable.

Data Availability Statements:

The original data sets are publicly available from the links provided in the paper. The synthetic data sets will be available from the corresponding author on reasonable request.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Elliott, M.; Houde, E.D.; Lamberth, S.J.; Lonsdale, J.A.; Tweedley, J.R. Management of Fishes and Fisheries in Estuaries. *Fish and Fisheries in Estuaries: A Global Perspective* **2022**, *2*, 706-797.
2. Zeller, D.; Pauly, D. Back to the future for fisheries, where will we choose to go? *Global Sustainability* **2019**, *2*, e11.
3. Lee, K.-H.; Noh, J.; Khim, J.S. The Blue Economy and the United Nations' sustainable development goals: Challenges and opportunities. *Environment international* **2020**, *137*, 105528.
4. Fasoulis, I. Governing the oceans: A study into Norway's ocean governance regime in the wake of United Nations Sustainable Development Goals. *Regional Studies in Marine Science* **2021**, *48*, 101983.
5. Obrecht, A.; Pham, M.; Spehn, E.; Payne, D.; Brémond, A.C.; Altermatt, F.; Fischer, M.; Passarello, C.; Moersberger, H.; Schelske, O. Achieving the SDGs with biodiversity. 2021.
6. Rowlands, G.; Brown, J.; Soule, B.; Boluda, P.T.; Rogers, A.D. Satellite surveillance of fishing vessel activity in the ascension island exclusive economic zone and marine protected area. *Marine Policy* **2019**, *101*, 39-50.
7. Gilman, E.; Kaiser, M.J.; Chaloupka, M. Do static and dynamic marine protected areas that restrict pelagic fishing achieve ecological objectives? *Ecosphere* **2019**, *10*, e02968.

8. ACTION, S.I. World Fisheries and Aquaculture. *Food and Agriculture Organization* **2020**, *2020*, 1-244.
9. Cowx, I.G.; Ogutu-Owhayo, R. Towards sustainable fisheries and aquaculture management in the African Great Lakes. *Fisheries Management and Ecology* **2019**, *26*, 397-405.
10. Diz, D. The ecosystem approach as a frame for SDG 14 implementation. *Ocean Yearbook Online* **2019**, *33*, 187-206.
11. Barbedo, J.G.A. A Review on the Use of Computer Vision and Artificial Intelligence for Fish Recognition, Monitoring, and Management. *Fishes* **2022**, *7*, 335.
12. Li, D.; Wang, Q.; Li, X.; Niu, M.; Wang, H.; Liu, C. Recent advances of machine vision technology in fish classification. *ICES Journal of Marine Science* **2022**, *79*, 263-284.
13. Alloghani, M.; Al-Jumeily, D.; Mustafina, J.; Hussain, A.; Aljaaf, A.J. A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised and unsupervised learning for data science* **2020**, 3-21.
14. Ogunlana, S.; Olabode, O.; Oluwadare, S.; Iwasokun, G. Fish classification using support vector machine. *African Journal of Computing & ICT* **2015**, *8*, 75-82.
15. Bimantoro, F.; Muhammad, D.A.; Marcellino, H.; Aranta, A. Classification of fish types using scale-invariant feature transform, bag of features and support vector machine. In Proceedings of the AIP Conference Proceedings, 2023; p. 040010.
16. Saberioon, M.; Císař, P.; Labbé, L.; Souček, P.; Pelissier, P.; Kerneis, T. Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (Oncorhynchus mykiss) classification using image-based features. *Sensors* **2018**, *18*, 1027.
17. Chhabra, H.S.; Srivastava, A.K.; Nijhawan, R. A hybrid deep learning approach for automatic fish classification. In Proceedings of the Proceedings of ICETIT 2019: Emerging Trends in Information Technology, 2020; pp. 427-436.
18. Pramunendar, R.A.; Wibirama, S.; Santosa, P.I. Fish classification based on underwater image interpolation and back-propagation neural network. In Proceedings of the 2019 5th International Conference on Science and Technology (ICST), 2019; pp. 1-6.
19. Almero, V.J.D.; Concepcion, R.S.; Sybingco, E.; Dadios, E.P. An image classifier for underwater fish detection using classification tree-artificial neural network hybrid. In Proceedings of the 2020 RIVF international conference on computing and communication technologies (RIVF), 2020; pp. 1-6.
20. Alsmadi, M.K.; Almarashdeh, I. A survey on fish classification techniques. *Journal of King Saud University - Computer and Information Sciences* **2022**, *34*, 1625-1638, doi:https://doi.org/10.1016/j.jksuci.2020.07.005.
21. Saitoh, T.; Shibata, T.; Miyazono, T. Feature points based fish image recognition. *International Journal of Computer Information Systems and Industrial Management Applications* **2016**, *8*, 12-22.
22. Kaharuddin, K.; Sholeha, E.W. Classification of Fish Species with Image Data Using K-Nearest Neighbor. *International Journal of Computer and Information System (IJCIS)* **2021**, *2*, 54-58.

23. Alsmadi, M.K.; Almarashdeh, I. A survey on fish classification techniques. *Journal of King Saud University-Computer and Information Sciences* **2022**, *34*, 1625-1638.
24. Hridayami, P.; Putra, I.K.G.D.; Wibawa, K.S. Fish species recognition using VGG16 deep convolutional neural network. *Journal of Computing Science and Engineering* **2019**, *13*, 124-130.
25. Montalbo, F.J.P.; Hernandez, A.A. Classification of fish species with augmented data using deep convolutional neural network. In Proceedings of the 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET), 2019; pp. 396-401.
26. Mol, J.J.; Jose, S.A. An Automated Fish Species Classification System Using Improved Alexnet Model. In Proceedings of the 2022 6th International Conference on Electronics, Communication and Aerospace Technology, 2022; pp. 121-126.
27. Bhanumathi, M.; Rithikab, R.; Roshni, R.; Selvarajb, S. Underwater Fish Species Classification Using Alexnet. **2022**.
28. Rauf, H.T.; Lali, M.I.U.; Zahoor, S.; Shah, S.Z.H.; Rehman, A.U.; Bukhari, S.A.C. Visual features based automated identification of fish species using deep convolutional neural networks. *Computers and electronics in agriculture* **2019**, *167*, 105075.
29. Agarwal, A.K.; Tiwari, R.G.; Khullar, V.; Kaushal, R.K. Transfer learning inspired fish species classification. In Proceedings of the 2021 8th International conference on signal processing and integrated networks (SPIN), 2021; pp. 1154-1159.
30. Mittal, S.; Srivastava, S.; Jayanth, J.P. A survey of deep learning techniques for underwater image classification. *IEEE Transactions on Neural Networks and Learning Systems* **2022**.
31. Zhang, S.; Liu, W.; Zhu, Y.; Han, W.; Huang, Y.; Li, J. Research on fish identification in tropical waters under unconstrained environment based on transfer learning. *Earth Science Informatics* **2022**, *15*, 1155-1166.
32. Ovalle, J.C.; Vilas, C.; Antelo, L.T. On the use of deep learning for fish species recognition and quantification on board fishing vessels. *Marine Policy* **2022**, *139*, 105015.
33. Yu, C.; Fan, X.; Hu, Z.; Xia, X.; Zhao, Y.; Li, R.; Bai, Y. Segmentation and measurement scheme for fish morphological features based on Mask R-CNN. *Information Processing in Agriculture* **2020**, *7*, 523-534.
34. Prasetyo, E.; Suciati, N.; Fatichah, C. A comparison of YOLO and mask R-CNN for segmenting head and tail of fish. In Proceedings of the 2020 4th International Conference on Informatics and Computational Sciences (ICICoS), 2020; pp. 1-6.
35. Conrady, C.R.; Er, Ş.; Attwood, C.G.; Roberson, L.A.; de Vos, L. Automated detection and classification of southern African Roman seabream using mask R-CNN. *Ecological Informatics* **2022**, *69*, 101593.
36. 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. *Limnology and Oceanography: Methods* **2016**, *14*, 570-585.
37. Deep, B.V.; Dash, R. Underwater fish species recognition using deep learning techniques. In Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks

(SPIN), 2019; pp. 665-669.

38. Fuchs, L.R.; Gällström, A.; Folkesson, J. Object recognition in forward looking sonar images using transfer learning. In Proceedings of the 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), 2018; pp. 1-6.
39. Mana, S.C.; Sasipraba, T. Analysis on Applying the Capabilities of Deep Learning Based Method for Underwater Fish Species Classification. In Proceedings of the Machine Learning and Big Data Analytics (Proceedings of International Conference on Machine Learning and Big Data Analytics (ICMLBDA) 2021), 2022; pp. 1-11.
40. Qiao, J. Framework for Fish Passage Design and Evaluation: Application for Emerald Shiners in Niagara River. State University of New York at Buffalo, 2019.
41. Pardede, J.; Sitohang, B.; Akbar, S.; Khodra, M.L. Implementation of transfer learning using VGG16 on fruit ripeness detection. *Int. J. Intell. Syst. Appl* **2021**, *13*, 52-61.
42. Belaid, O.N.; Loudini, M. Classification of brain tumor by combination of pre-trained vgg16 cnn. *Journal of Information Technology Management* **2020**, *12*, 13-25.
43. Majid, A.; Khan, M.A.; Yasmin, M.; Rehman, A.; Yousafzai, A.; Tariq, U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. *Microscopy research and technique* **2020**, *83*, 562-576.
44. Younis, A.; Qiang, L.; Nyatega, C.O.; Adamu, M.J.; Kawuwa, H.B. Brain tumor analysis using deep learning and VGG-16 ensembling learning approaches. *Applied Sciences* **2022**, *12*, 7282.
45. Kaur, T.; Gandhi, T.K. Automated brain image classification based on VGG-16 and transfer learning. In Proceedings of the 2019 International Conference on Information Technology (ICIT), 2019; pp. 94-98.
46. Zhao, D.; Zhu, D.; Lu, J.; Luo, Y.; Zhang, G. Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16. *Symmetry* **2018**, *10*, 519.
47. *FishBse, https://www.fishbase.org.au/v4(accessed on 16 March 2023)*.
48. *ImageNet, https://www.image-net.org/(accessed on 15 March 2023)*.
49. Perez, L.; Wang, J. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621* **2017**.
50. Mikołajczyk, A.; Grochowski, M. Data augmentation for improving deep learning in image classification problem. In Proceedings of the 2018 international interdisciplinary PhD workshop (IIPhDW), 2018; pp. 117-122.
51. Chua, L.O.; Roska, T. The CNN paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **1993**, *40*, 147-156.
52. Bhatt, D.; Patel, C.; Talsania, H.; Patel, J.; Vaghela, R.; Pandya, S.; Modi, K.; Ghayvat, H. CNN variants for computer vision: history, architecture, application, challenges and future scope. *Electronics* **2021**, *10*, 2470.
53. Fitrianah, D.; Suryaningrum, K.M.; Sagala, N.T.M.; Ayumi, V.; Lim, S.M. Fine-Tuned MobileNetV2 and VGG16 Algorithm for Fish Image Classification. In Proceedings of the 2022 International Conference

on Informatics, Multimedia, Cyber and Information System (ICIMCIS), 2022; pp. 384-389.
54. Meslet-Millet, F.; Chaput, E.; Mouysset, S. SPPNet: An approach for real-time encrypted traffic classification using deep learning. In Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM), 2021; pp. 1-6.
55. Chen, X.; Gupta, A. An implementation of faster rcnn with study for region sampling. *arXiv preprint arXiv:1702.02138* **2017**.
56. Qassim, H.; Verma, A.; Feinzimer, D. Compressed residual-VGG16 CNN model for big data places image recognition. In Proceedings of the 2018 IEEE 8th annual computing and communication workshop and conference (CCWC), 2018; pp. 169-175.
57. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence* **2015**, *37*, 1904-1916.
58. Luque, A.; Carrasco, A.; Martín, A.; de Las Heras, A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition* **2019**, *91*, 216-231.

Figures

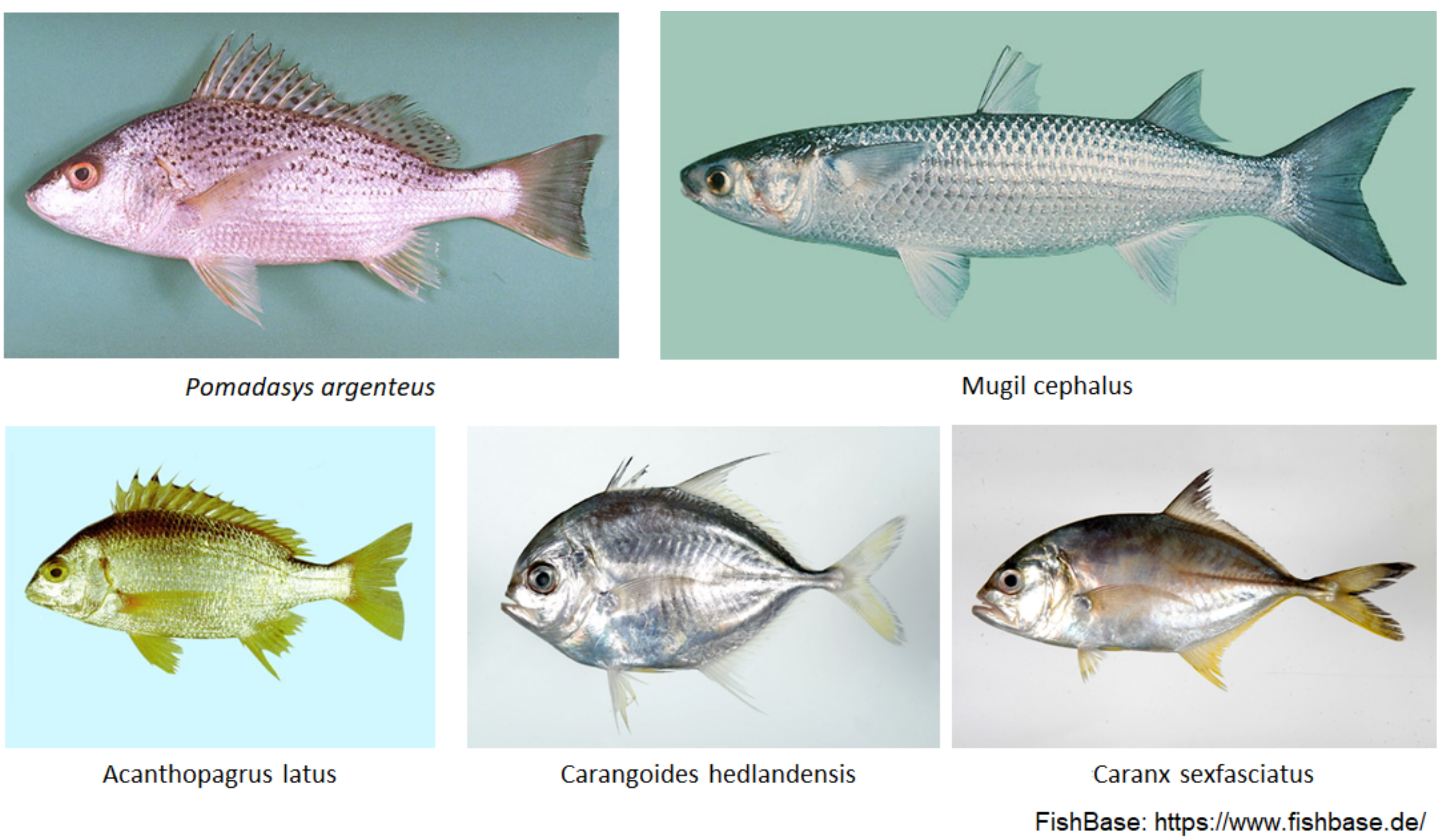

Figure 1

Fish species selected for classification

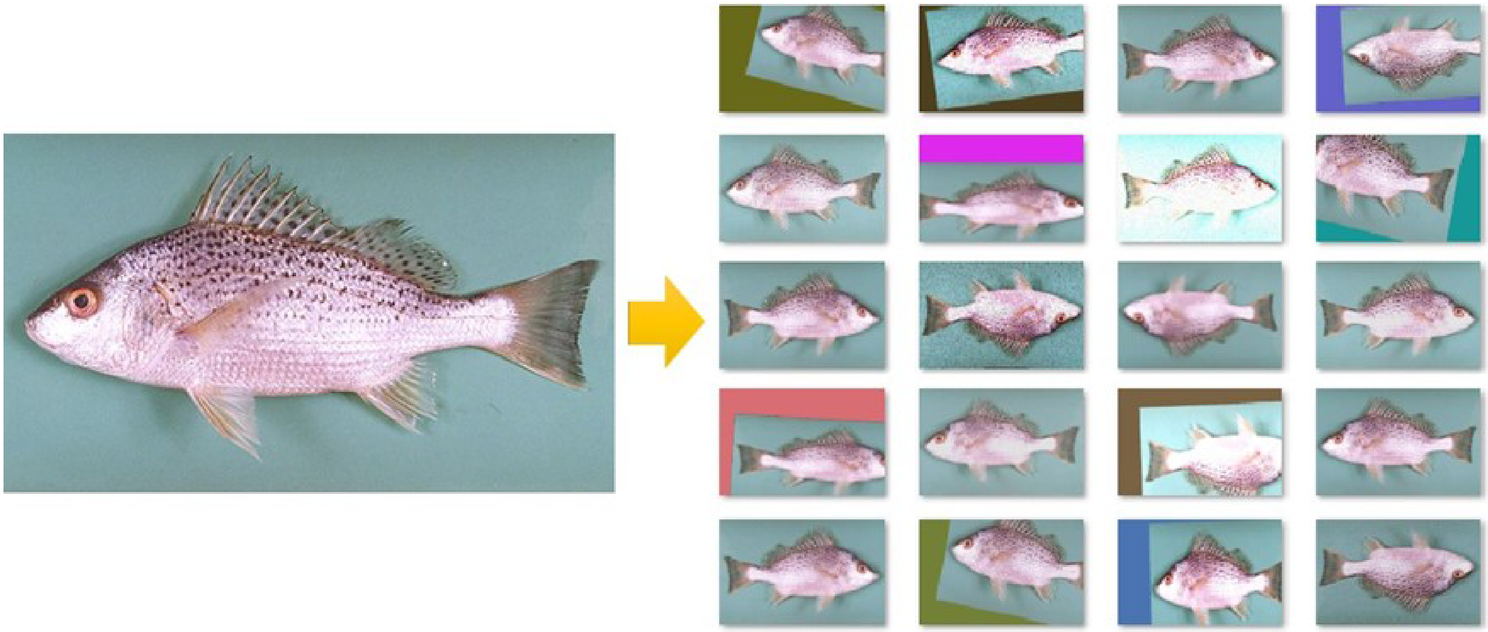

Figure 2

Augmentation for raw image

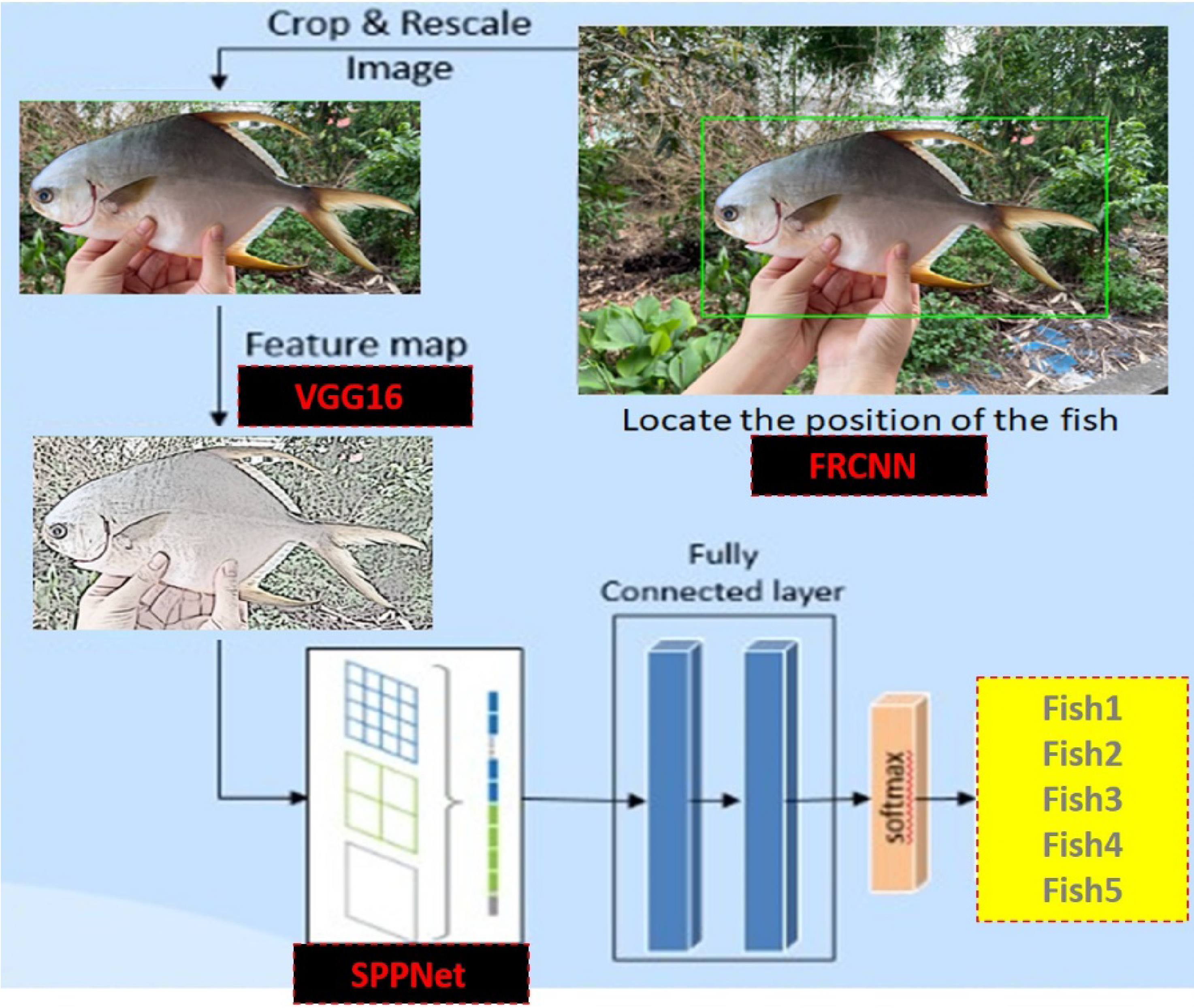

Figure 3

Architecture of proposed FRCNN-VGG16-SPPNet model

		Actual Class	
		Positive(P)	Negative(N)
Predicted Class	Positive(P)	True Positive(TP)	False Positive (FP)
	Negative(N)	False Negative (FN)	True Negative(TN)

Figure 4

Confusion matrix for binary classification

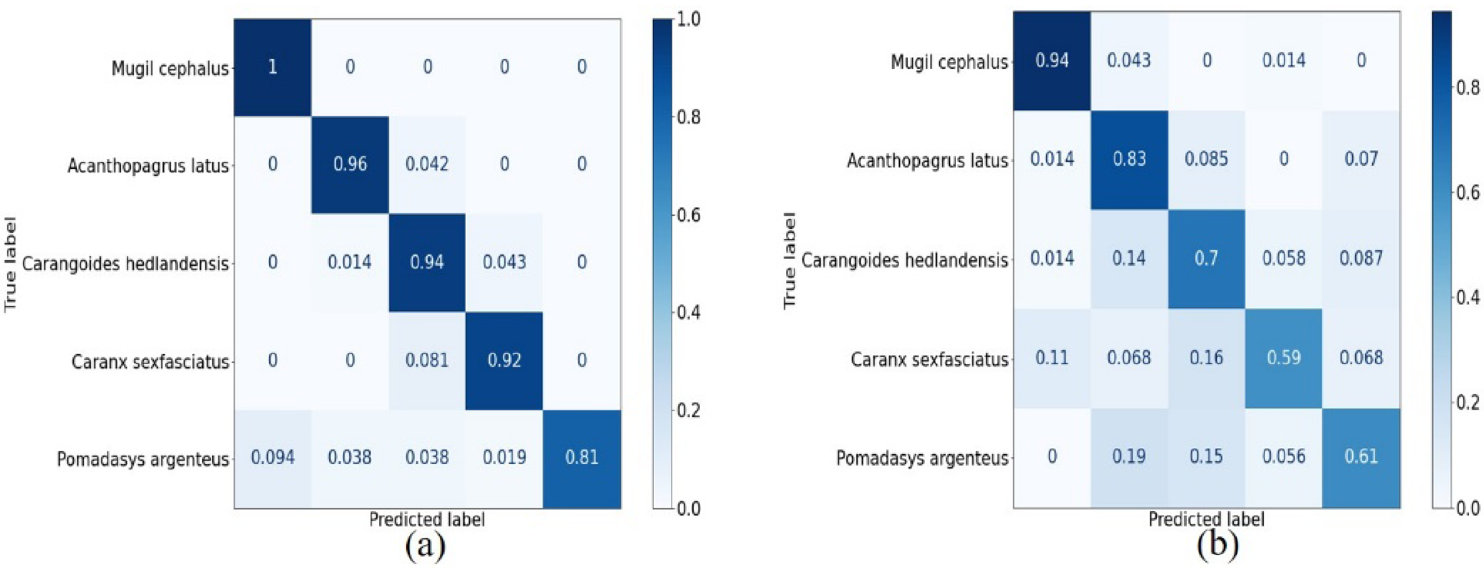

Figure 5

Confusion Matrix for FRCNN-VGG16-SPPNet (a) and single VGG16 model (b)

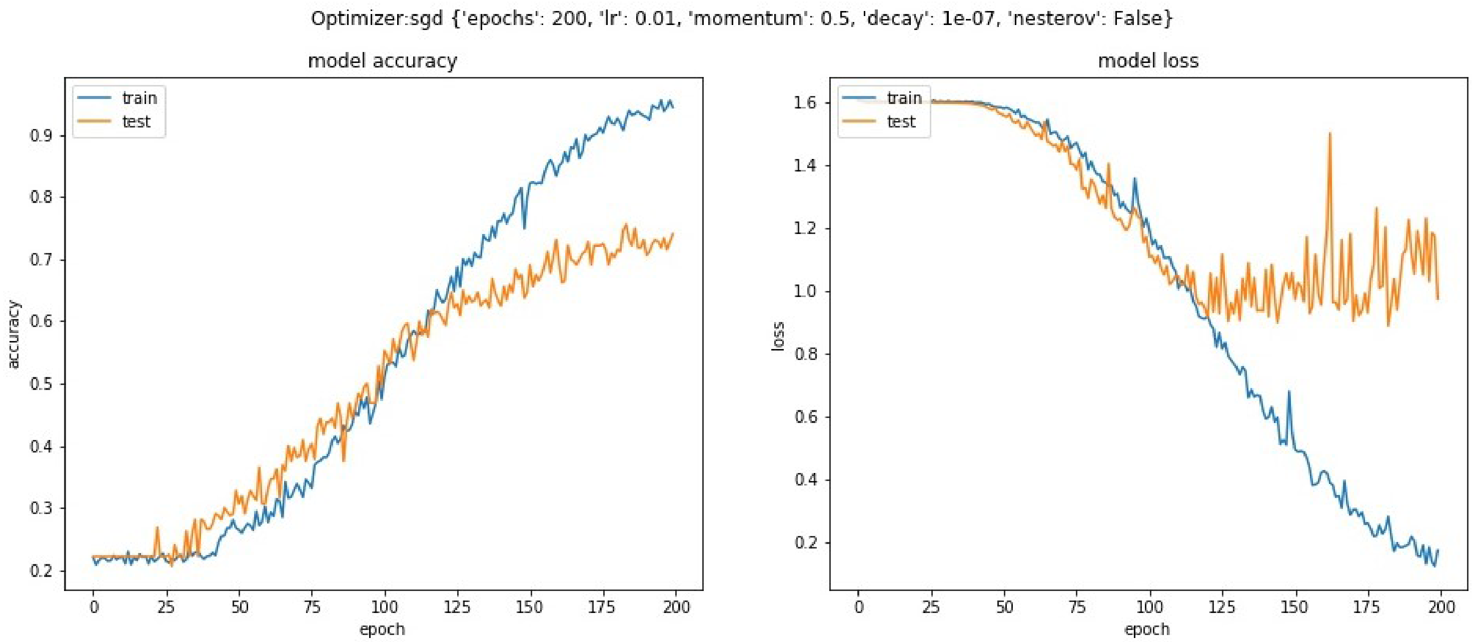

Figure 6

Learning curves for single VGG16 model

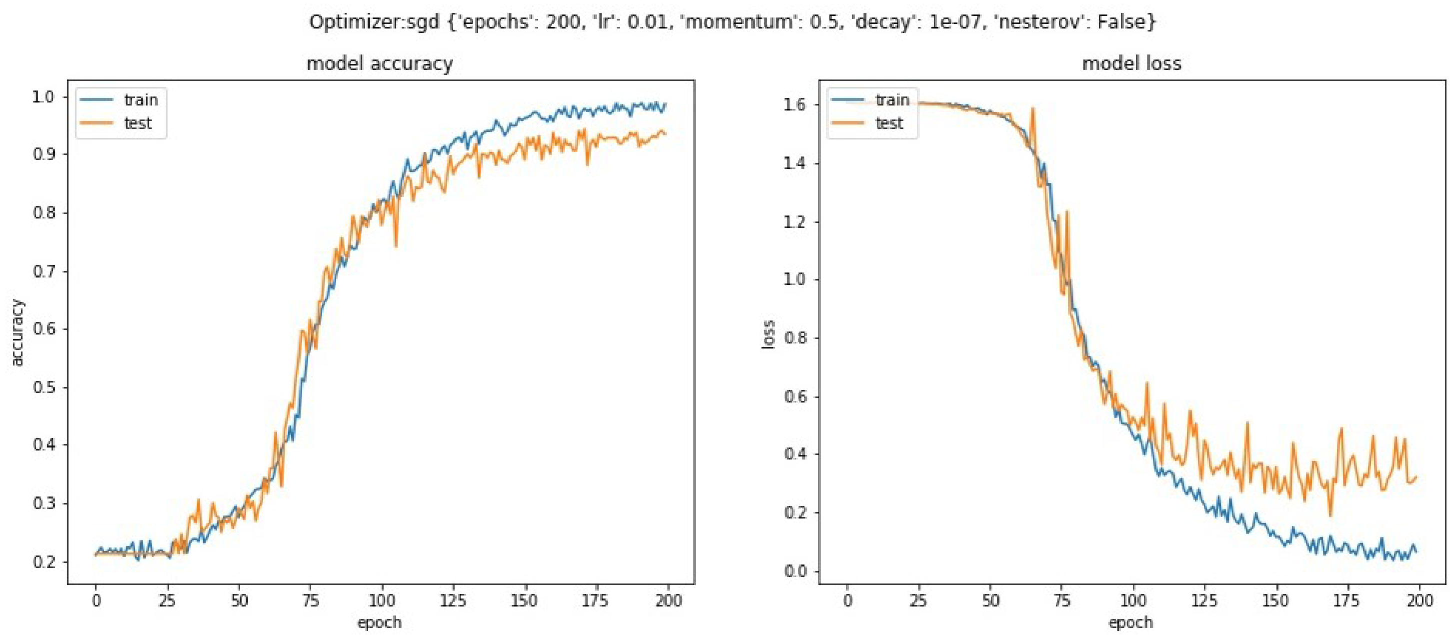

Figure 7

Learning curves for the proposed FRCNN-VGG16-SPPNet model