PAPER

Extreme Kernel Machine (EKM)-based slim network for accurate fish species recognition to mitigate its extinction

To cite this article: Anantha Prabha P *et al* 2025 *Eng. Res. Express* **7** 035250

IOP Publishing Eng. Res. Express 7 (2025) 035250 https://doi.org/10.1088/2631-8695/adf279

Engineering Research Express

RECEIVED
17 March 2025

REVISED
21 June 2025

ACCEPTED FOR PUBLICATION
21 July 2025

PUBLISHED
29 August 2025

PAPER

Extreme Kernel Machine (EKM)-based slim network for accurate fish species recognition to mitigate its extinction

Anantha Prabha P[1,*], **Sengathir J**[2], **Deva Priya M**[3] **and Suguna Devi R**[4]

[1] Department of Computer Science and Engineering, PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India
[2] Department of Information Technology, CVR College of Engineering, Hyderabad, Telangana, India
[3] Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India
[4] Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, Tamil Nadu, India
[*] Author to whom any correspondence should be addressed.

E-mail: ap.prabha@gmail.com, j.sengathir@gmail.com, meetdevapriya@gmail.com and rsdsaran2008@gmail.com

Keywords: fish extinction, extreme kernel machine, extreme learning machine, slim pre-trained network, transfer learning, deep learning

Abstract

Around the globe, several fish species are on the verge of extinction due to overfishing and environmental degradation. An efficient automated fish monitoring system is crucial for rapid and accurate identification of fish species in water bodies. Deep Learning (DL)-based fish species recognition methodologies are highly suitable, as they offer efficient image classification with improved accuracy compared to traditional methods. However, these models are prone to overfitting, and there are chances for the performance of classifiers to be skewed toward common fish species due to the limited availability of samples. To address these issues, in this paper, a Slim Pre-Trained Network with Extreme Kernel Machine (SPTN-EKM)-based fish recognition mechanism is proposed. It focuses on integrating the benefits of SPTN with EKM such that precise classification is achieved. The proposed SPTN-EKM model contextually leverages the potentialities of optimized SPTN, extracts high-level features, and effectively classifies fish species using the EKM classifier. The experimental investigation of the proposed SPTN-EKM and the baseline fish species recognition model is conducted using two publicly available datasets, namely, Fish-Pak and QUT datasets, and a synthetic dataset called the Cephalopod dataset, which includes images of fish species found in the Kasimedu and Tuticorin coastal areas of the Bay of Bengal. The experimental outcomes of the proposed SPTN-EKM model with respect to Fish-Pak, QUT, and Cephalopod datasets confirm an increased accuracy of 98.61%, 96.34% and 93.34% respectively, for samples considered for recognition. It is also identified to significantly minimize the training time compared to the baseline DL-based models considered for fish species recognition.

1. Introduction

A significant part of the human population consumes fish as it is the prime source of protein, while another sector has fishing as its main source of living (FAO 2012, Purcell *et al* 2013). Fishing plays a dominant role in triggering a decrease in the marine mammal population. Vessels involved in fishing lead to faster depletion of species than replenishment of stocks. Overfishing seems to be risky as a particular species of fish may be caught at an increased rate faster than the rate of replenishment, resulting in their diminution. This may be seen in all water bodies, including rivers, lakes, ponds, and oceans, instigating exhaustion of resources, reduction of natural growth rates and biomass levels, thereby leading to unsustainability of fish populations (Thilsted *et al* 2016). Overfishing of particular species, like sharks, leads to instability of the entire marine ecosystem. It is associated with bycatch or catching of unwanted sea life while hunting for a different species. This may lead to depletion of fish, including sea turtles and cetaceans (Zaneveld *et al* 2016). The consequences may extend beyond the maritime ecosystem. Owing to overfishing in the seas as well as environmental deterioration, there are high chances of fish species to become extinct. To handle this issue, an automated system that is efficient in monitoring fish for quick identification is mandatory. DL-based systems for identifying fish species are becoming popular. Due to limited fish sample size, the model may overfit, causing the classifier to bias towards species that are more frequently seen in the samples.

Identifying fish species and monitoring their abundance is crucial for addressing issues related to their extinction and bycatch. The species can be recognized based on their biochemical and molecular taxonomic features (Quinteiro *et al* 1998; Mackie *et al* 1999). Presently, there is a great demand for taxonomic proficiency, and there is a lack of efficient taxonomists. Experts deprived of prior knowledge about taxonomy, including fishery investigators, inspectors, customs officials, dealers, researchers, etc, may be involved in the challenging task of identifying aquatic species due to the availability of imprecise and inadequate information concerning the encountered species. Products like FishFinder Program's species catalogues and field guides of the Food and Agriculture Organisation (FAO) aid in finding the regions of availability, whereas digital tools like FishBase1 and Catalog of Fishes2 assist in determining the scientific name of the species (Fischer 2014). However, it requires a significant amount of effort to ensure that aquatic resources are accurately identified. Implanting light-weight networks with improved classification accuracy into smart monitoring devices enables effective identification of fish categories.

The main aim of this work is to design such networks that offer improved accuracy, supporting efficient feature extraction and classification. Existing works use Convolutional Neural Network (CNN) like AlexNet, VGG, etc, for classification, which may not offer better performance and involve an increased training period. Based on the above-mentioned facts, a novel mechanism is proposed to handle these issues. A hybrid scheme called Slim Pre-trained Network with External Kernel Machine (SPTN-EKM) that deals with fish species categorization is proposed. It specifically uses the merits of optimized SPTN and EKM-classifier such that accurate recognition of fish species is achieved with maximized accuracy and precision, independent of the publicly available and synthetic datasets created for experimental purposes. This DL-based model plays an indispensable role in deriving high-quality and reliable data features from the input images, such that accurate classification can be achieved during the process of recognition.

The paper is organised as follows. Section 2 reviews the existing methodologies and research efforts in fish species identification, emphasizing the role of DL and hybrid models. Section 3 details the datasets utilized in this research. Section 4 outlines the proposed system. Section 5 covers the experimental results and their analysis, while section 6 gives the conclusion.

2. Related work

Technical advancements are rapidly being employed in all kinds of fisheries to enhance the identification of fish species. Computer vision-dependent technologies are highly used in studies related to marine ecology to lessen the amount of human processing and improve classification accuracy (Li *et al* 2015, Qin *et al* 2016, Jin and Liang 2017).

Salman *et al* (2016) have presented a DL-based scheme that is appropriate for the classification of fish species. It employs the CNN model to support hierarchical integration of features for gaining information about the visual characteristics of fish species. It is found to be robust to environmental, intra-species and inter-species variabilities. The scheme evades the necessity for obviously mining features from fish images using numerous image processing schemes for LifeCLEF14 and LifeCLEF15 standard fish datasets. This CNN-based scheme offers a better classification rate in contrast to the same and cross-dataset training-testing protocols. Yet, its dependence on huge datasets and computational resources poses challenges for real-time applications on resource-constrained devices and generalizability across different aquatic ecosystems. Ishaq *et al* (2017) have used DL for precise categorisation of whole-body zebrafish distortions in multi-fish microwell plates. The proposed scheme is efficient in handling raw images without demanding expert knowledge for the optimisation of segmentation factors. However, it focuses on a specific species, which may not generalize well to other fish species with different morphological features. Additionally, the method requires high-quality imaging and sufficient data to efficiently train the model, which may not always be feasible in real-world scenarios. Siddiqui *et al* (2018) have proposed an automatic fish species classification method for underwater videos by leveraging pre-trained Deep Neural Network (DNN) models to address the challenge of limited labelled data. The approach demonstrates the potential to classify fish species effectively with minimal annotation by using Transfer Learning (TL) techniques. However, the method depends on pre-trained models, which may not always capture the unique features of specific underwater environments.

Liu *et al* (2019) have developed a real-time marine animal image classification system that employs MobileNet for embedded systems. The approach demonstrates the feasibility of deploying an efficient, lightweight DL-based classification model for resource-constrained devices, making it suitable for real-time applications. However, its limitations include potential challenges in adapting to diverse marine environments with varying species and image qualities. Additionally, reliance on TL may result in suboptimal performance in case the pre-trained models do not align well with the specific characteristics of marine animal images. Park *et al* (2019) have applied CNNs for fish species classification, leveraging CNNs' ability to extract intricate features from aquatic images. This approach demonstrates high classification accuracy and robustness in identifying various fish species. The study emphasizes the efficiency of CNNs in handling large and diverse datasets, making them suitable for real-world applications. However, the model is dependent on high-quality, labelled image datasets, which may not always be readily available in marine environments. Mathur *et al* (2020) have proposed a TL-based classification model using Artificial Neural Network (ANN). The restricted availability of images makes it challenging to train the networks as they necessitate massive datasets. To decrease the demand for training data, an algorithm involving the pooling of cross-convolutional layer on a pre-trained CNN is proposed. The work emphasizes on development of an automatic system for classification that is efficient in identifying and categorizing fishes from underwater images obtained from videos. It considers a dataset with 27,370 fish images for assessment. The proposed scheme seems to be an effective substitute to manual identification by experts and is beneficial for observing fish biodiversity in natural habitats. However, the generalizability of the model may be constrained when applied to species with significant variations

Identification of rare fish species is challenging owing to the reduced number of fish samples and an imbalanced number of categories. The model may be over-fitted, and the classifier's performance may be biased towards fish species in several samples. To handle these problems, Xu *et al* (2021) have proposed a scheme for identifying fish species using SE-ResNet152 along with class-stable focal loss. Initially, the scheme involves a visualization study and pre-processing of images of fish datasets. Secondly, it builds an SE-ResNet152 model for extracting features and applies it to the dataset. It uses class-stable focal loss to train the SE-ResNet152 model and identifies fish species from head, body, and scale views of fishes. The proposed scheme applied to Fish-Pak public dataset offers improved accuracy. However, the approach demands high computational resources, particularly in fine-tuning pre-trained models. Additionally, performance may degrade when applied to datasets with significant variability in environmental conditions or fish appearance. Zhang *et al* (2022) have proposed a TL-based fish detection scheme for fishes in tropical waters under an unconstrained setting. It builds RestNet50-based CNN to associate fish recognition before and after the application of TL. This TL-based system offers better accuracy and less loss by using a trained ImageNet model. The indicators converge when the model is trained for 150 epochs. TL models may have difficulty in generalizing across various species or geographic regions, which can limit their effectiveness in real-world applications. Jiang *et al* (2022) have developed a fish recognition method tailored for complex underwater scenes using targeted sample TL. This scheme enhances recognition accuracy by focusing on transferring knowledge from relevant datasets and addressing challenges like occlusion, low visibility, and complex backgrounds. The study highlights the effectiveness of sample-specific TL in improving model performance under challenging underwater conditions. However, a limitation of this method is its reliance on carefully curated and domain-specific datasets, which may not be available.

Shang *et al* (2023) have proposed an improved ResNet-RS model tailored for underwater fish image classification. The approach incorporates instance standardization along with a batch normalization layer to enhance generalization capabilities and robustness. The model uses a pre-activated residual network to improve feature extraction and adopts the Lion optimizer to accelerate convergence rates. Simulation results demonstrate superior classification accuracy in contrast to traditional methods, thus highlighting the effectiveness of integrated architectural enhancements and advanced optimization techniques for underwater scenarios. A limitation of the improved ResNet-RS model is its high demand for computational resources, which may restrict its use in environments with limited resources. Satoto *et al* (2023) have presented a marine fish species classification approach using TL combined with a Residual Network (ResNet). This method effectively leverages pre-trained models to improve feature extraction and classification accuracy for diverse marine fish species. The study highlights the efficiency of TL in reducing training time while maintaining robust performance. However, this method can be applied only to labelled and high-quality images. Malik *et al* (2023) have proposed a multi-classification DNN for identifying fish species using camera-captured images. This method uses DL architectures to effectively handle diverse fish species in natural environments. The study offers better classification accuracy and robustness across various datasets. Nevertheless, it involves increased computational overhead for training DNN, particularly when applied to large-scale or high-resolution image datasets.

Gao *et al* (2024) have proposed an optimized ResNet50 model to address challenges in marine fish recognition, such as complex underwater environments and species variability. This approach focuses on refining the ResNet50 structure to improve feature extraction as well as classification accuracy. By integrating optimization techniques and adapting the model for underwater conditions, it achieves notable improvements in both performance and robustness. However, the approach has limitations, including its reliance on large, high-quality datasets, which may not be readily available for all species. Sirigineedi *et al* (2024) have proposed a hybrid VGG16-Darknet architecture focusing on the enhancement of fish image detection accuracy and speed.

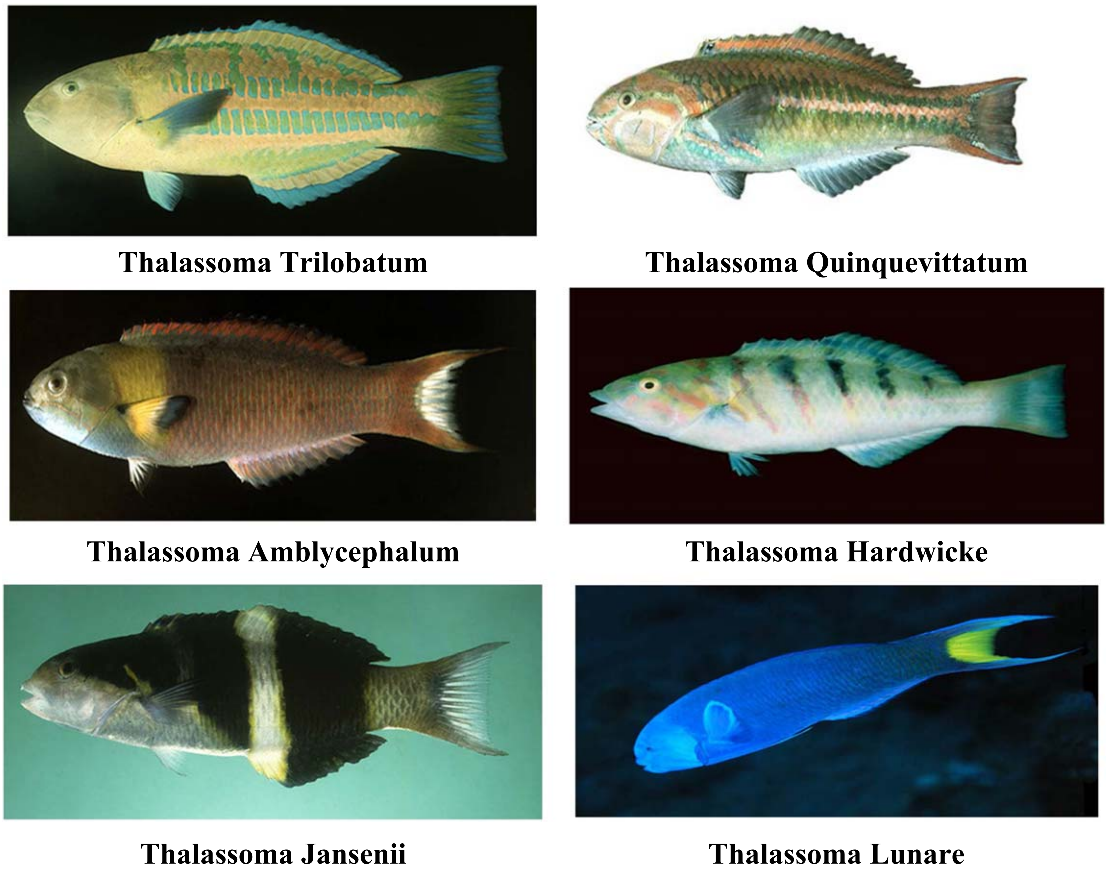


Figure 1. Sample fish species of the QUT dataset (Controlled). Reproduced from [Anantharajah *et al* 2014].

This study highlights the effectiveness of CNNs in feature extraction and object detection, offering improved performance in terms of accuracy and processing efficiency. It can be applied in real-time and is scalable. However, this method is not efficient in handling complex underwater environments with varying lighting conditions and occlusions. Ataç and Şahin (2024) have evaluated the performance of various TL techniques in classifying fish species. This study compares different pre-trained models, highlighting the effectiveness in improving the classification accuracy of diverse fish datasets. The findings emphasize the advantages of TL in minimizing training time and computational resources. However, the model is sensitive to variations in image quality and environmental conditions, which can affect classification performance.

3. Image datasets

In the present experimental studies, two publicly available datasets namely, QUT and Fish-Pak dataset and a tailored synthetic Cephalopod dataset are used. A Detailed description of datasets is given below.

3.1. QUT dataset

Around 3960 images in the QUT fish collection (Anantharajah *et al* 2014, QCR.CAI 2016 https://wiki.qut.edu.au/display/cyphy/Fish+Dataset) are taken under 'Controlled,' 'out-of-water' and '*in situ*' settings.

The 'controlled' environment includes photographs of a variety of fish species taken against a consistent background. Photographs in 'out-of-the-water' are captured without altering the background and under extremely limited lighting. The photographs in '*in situ*' category are captured undersea from a natural habitat. In this research, six fish species were chosen and photographed under diverse conditions. Figure 1 shows some sample images from the dataset.

3.2. Fish-Pak dataset

Other images utilised in the experiment are from the Fish-Pak Version 3 dataset (Shah *et al* 2019, Rauf 2019 https://data.mendeley.com/datasets/n3ydw29sbz/1). The images in this dataset are captured using a Canon EOS 1300D at a resolution of 5184 $\times$ 3456 from 3 separate locations in Gujarat, Punjab, Pakistan: Head Qadirabad, Head Marala and Chenab River. Figure 2 shows some instances of fish species from Fish-Pak dataset.

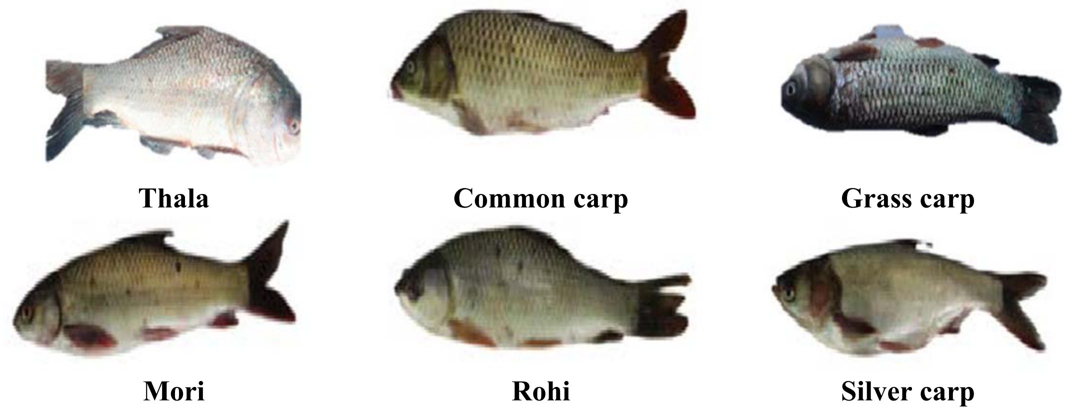


Figure 2. Sample fish species of the Fish-Pak dataset. Reproduced from [Shah *et al* 2019].

Table 1. Details of the Fish-Pak dataset.

Classes	Body	Head	Scale	Total
Catla catla (Thala)	20	25	11	56
Cyprinus carpio	50	64	44	158
Ctenopharyngodon idella	11	16	9	36
Cirrhinus mrigala	70	100	71	241
Labeo rohita	73	114	62	249
Hypophthalmichthys molitrix (Silver carp)	47	71	57	175
Total	271	390	254	915

Table 1 gives an overview of the Fish-Pak dataset. 271 body images out of 915 total images of three different views (body, head and scale) are used in the present study.

3.3. Cephalopod dataset

This dataset comprises Cephalopod images captured using a Canon EOS 1000D camera with a resolution of 10.1 megapixels (3888 $\times$ 2592). These images are taken from the Kasimedu and Tuticorin coastal regions along the coasts of the Bay of Bengal. The dataset includes six prevalent Cephalopod species namely, Sepia prashadi, Sepia pharaonis, Sepia aculeata, Sepioteuthis lessoniana, Uroteuthis duvauceli and Doryteuthis sibogae. Figures 3(a) and (b) (Prabha *et al* 2023) show the field and sample images. Cephalopod species are widespread across India and are known for their high reproductive rates.

These species play a crucial role in classification studies as they represent distinct classes within the image dataset involved in training and testing of DL models. Their presence helps the model to identify and differentiate categories based on specific visual characteristics, thus enhancing the system's accuracy, reliability and ability to generalize to new data. A total of 492 images were collected from coastal regions with the following distribution: Doryteuthis Sibogae (84), Sepia Aculeata (88), Sepia Pharaonis (80), Sepia Prashadi (88), Sepioteuthis Lessoniana (78), and Uroteuthis Duvauceli (74).

4. Image augmentation and scaling

Data augmentation is done to overcome overfitting problems by increasing the size of the fish dataset (Liu *et al* 2020). To increase the size of the dataset, a few transformation schemes like the following are used:

- Random rotation of images within a particular range using an unchanging distribution
- Displacement of random images in both vertical and horizontal directions

A portion of the images is horizontally flipped, randomly scaled with some shearing modifications. The image data generator (Montserrat *et al* 2017) considers a set of images and arbitrarily changes them into a fresh set. The original dataset comprises 492 images. To increase the dataset size, and improve performance and

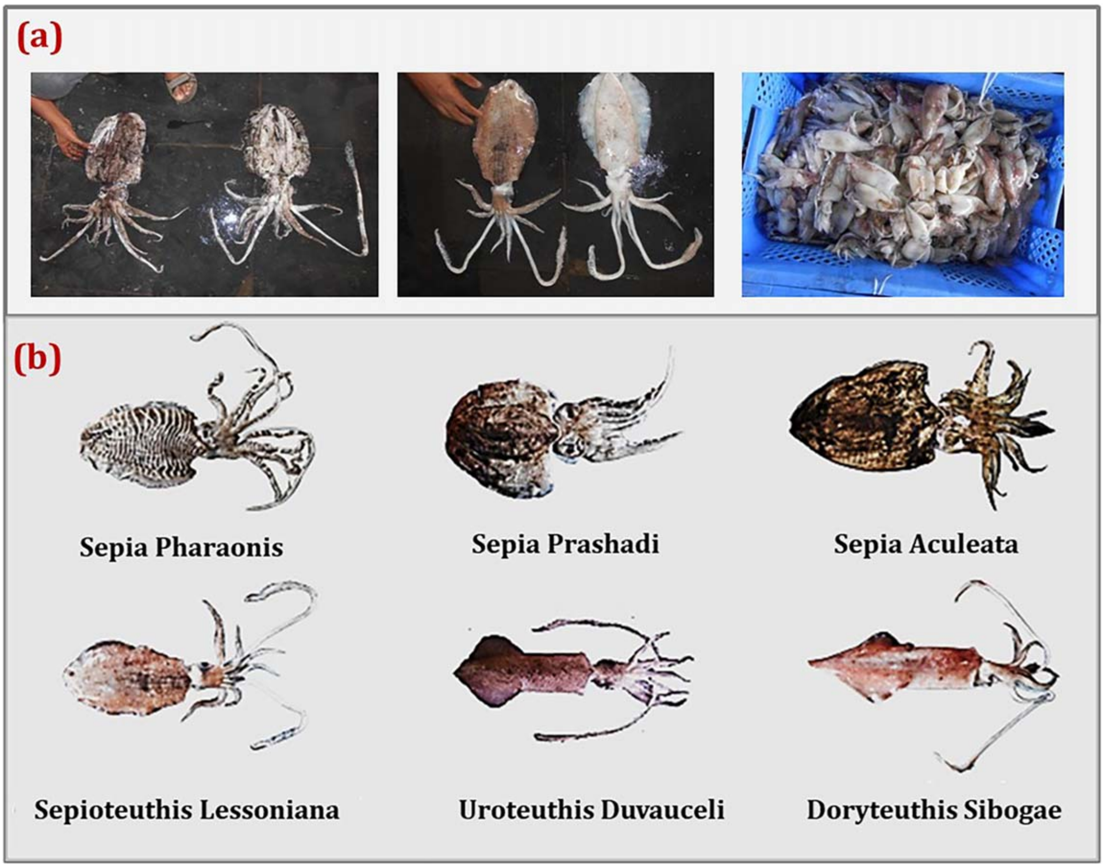


Figure 3. (a): Field images of Cephalopod species. (b) Sample images of Cephalopod species. Reproduced from [Prabha *et al* 2023]. CC BY 4.0.

generalization capability of the model, data augmentation techniques such as rotation, flipping, scaling and shearing are applied. Each image is augmented 7 times, resulting in 3,444 new images. Including original images, the final dataset contains a total of 3,936 images. The images are scaled to a size of 224 × 224 pixels before applying it to any network.

5. Proposed Slim Pre-trained Networks with External Kernel Machine (SPTN-EKM)-based fish classifier

A hybrid scheme called SPTN-EKM is proposed for identifying the fish species. It is efficient in handling challenges related to insufficient data and increased training time. The proposed scheme involves two phases namely, feature extraction using a pre-trained network and classification of images using EKM. The system architecture of the proposed scheme is shown in figure 4.

5.1. Feature extraction using SPTN

Deep CNN (DCNN) is used for building a thin model which supports devices with reduced specifications, including mobile hardware, embedded and microcomputers. Thin models involve less computational complexity, thus reducing processing time (Winoto *et al* 2020). The main goal of the work is to build a simple framework, demanding reduced training time. Hence, light-weight pre-trained models are used in fish classification.

The network models listed in table 2 are of size less than 100 MB. In this work, pre-trained network models like MobileNetV2, NASNetMobile, DenseNet201, Xception, InceptionV3 and ResNet50 are used for determining features from images. For performance evaluation, an additional Softmax layer is appended to each pre-trained network to support classification. This Softmax layer is trained for 30 epochs using the Adam optimizer. The extracted features are fed into external classifiers, including Extreme Learning Machine (ELM) and the proposed EKM for comparative analysis.

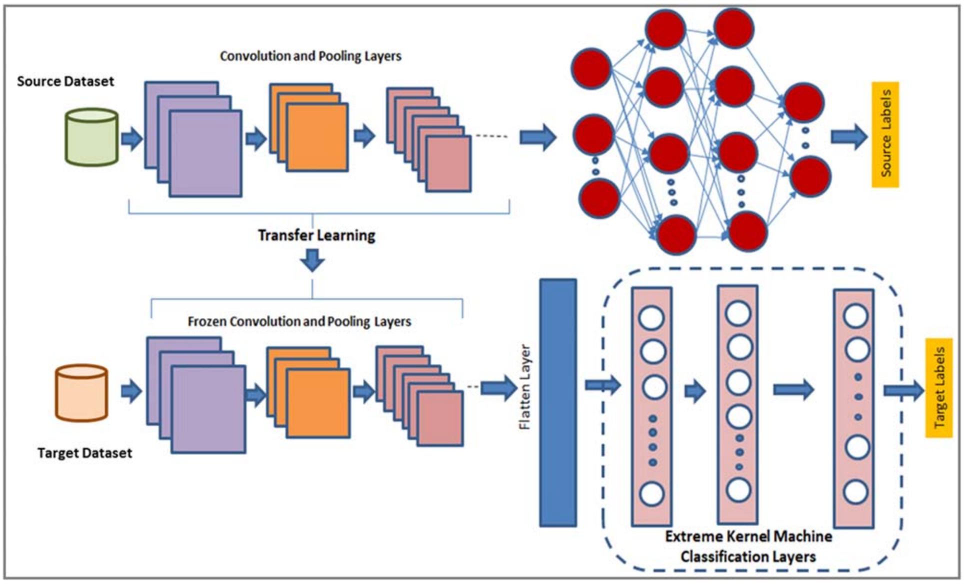


Figure 4. Proposed SPTN-EKM architecture for fish recognition.

Table 2. Sizes of models.

Model	Size (MB)
MobileNetV2	16
NASNet Mobile	23
DenseNet201	80
Xception	88
InceptionV3	92
ResNet50	98

5.2. EKM-based classification

ELMs are Feed Forward Neural Networks (FFNNs) designed for tasks such as classification, clustering, regression, sparse approximation, feature learning and compression (Huang *et al* 2011). They usually consist of a single hidden layer, where the parameters of the hidden nodes comprising weights and biases are arbitrarily assigned and remain fixed during training. This approach eliminates the need for iterative tuning of these parameters, thus simplifying the training process. The output weights of the hidden nodes are typically determined through a linear model, specifically the least-squares method.

Compared to backpropagation-trained networks, these models are efficient in producing efficient generalisation and learning at an increased rate. The ELM algorithm does not necessitate tuning of hidden nodes/neurons. It creates biases and input weights for hidden layers, distributes hidden nodes at random, and uses least squares methods to determine output weights. This demonstrates the quick computation speed of ELM, thus making it the preferred choice.

The output of the second (hidden) layer is generated by randomly initialising the bias and weights of the input layer. In ELMs, the parameters are randomly initialized and remain fixed during training. Random initialization simplifies the training process by eliminating the need for iterative tuning of parameters. Consequently, ELMs involve less training time and are less susceptible to overfitting compared to traditional neural networks that require extensive parameter optimization. As a result, ELMs are frequently used for classification. By using linear combinations of randomly mapped values, the appropriate class is determined from the output layer. EKM maps features into infinite dimensions that can effectively distinguish features.

Single-layered FFNs are a kind of ELMs which modify feature matrix'$[f_1, f_2, \ldots, f_n] \in E_d \times n$' by using matrix multiplications in the hidden layer with 'h' neurons '$M_i \in E^{h \times d}$', where 'E_d' represents a vector space over a field 'd'. The variable 'n' signifies the quantity of features or data points in matrix 'F'. 'M_i' denotes the weight matrix and '$E^{h \times d}$' represents the vector space over a field 'E'.

Input features can be arbitrarily rotated as every square sub-matrix is of 'd' dimensions in 'M_i'. This enables transformation of input features from 'E^d' to ELM-space 'E^h'. Non-linear function (f(A)) represented as an activation function is capable of managing the output of hidden neurons. Rectified Linear Unit (ReLU) is used for simplifying computation as shown in equation (1).

$$f(A) = \max\{0, A\} = \begin{cases} 0, & \text{if } A_{i,j} < 0 \\ A_{i,j}, & \textit{Else} \end{cases} \tag{1}$$

The output of the activation function is given by $O \equiv f(M_i.F)$

Lastly, the product of the hidden-to-output-layer matrix $(M_o \in E^{c\times h})$ and output $(O \in E^{h\times n})$ is computed, where 'c' represents the number of classes in the dataset. The solution of optimisation is defined in the matrix as shown in equation (2).

$$\min_{M_o \in E^{c}\times h} \frac{1}{2} M_o.O - Y^2 \tag{2}$$

One-hot label encoding of input feature matrix 'F' is represented as '$Y \in E^{c\times n}$'. The main part of the ELM classification algorithm is shown in equation (3). The set of scores assigned by ELM to every input feature vector of each class is denoted as '$Y \in E^{c\times n}$'.

$$\hat{Y} = M_o.O = M_o.f(M_i.F) \tag{3}$$

Classification is done by assigning a feature vector to the class with the highest score.

Optimisation constraints can be included in equation (2) to enhance the performance of the model. 'L_2' regularisation is included in equation (4).

$$\min_{M_o \in E^{-}\times h} \frac{1}{2} \|M_o.O - Y\|^2 + \frac{\varphi}{2} \|M_o\|^2 \tag{4}$$

Parameter 'φ' focuses on monitoring the level of regularisation that has control on the model. The main aim of applying regularisation is to improve the model's generality by reducing the quantity of time spent on overfitting the training set.

The solution for 'M_o' is given in equation (5), where 'I' represents the Identity matrix of suitable dimension. The output of the activation function (O) and 'M_i' remain unaffected.

$$M_o = Y(O^T.O + \lambda.I)^{-1}O^T \tag{5}$$

ELM facilitates easy computation of output weights combined with the kernel's similarity measure. Perception specifies that mining more amount of information from the training set enhances the classification performance of the model. This inclusion does not change the formulation of the optimisation problem, but modifies the included matrices.

Though ELM offers fast learning and generalization, it may be tedious to handle complex and non-linearly separable data. To address this, Extreme Kernel Machines (EKMs) (Karlsson and Rosvall 2017) extend ELMs by implicitly mapping input features into a high-dimensional kernel space without explicitly determining hidden layer activations. The inner product in this feature space is represented using a kernel function.

Instead of matrix 'O' in equation (4), a Kernel matrix (K_t) is introduced. As optimisation is handled in the training stage, the kernel function applied on 'i^{th}' and 'j^{th}' vectors in the training set are represented as 'K_t, $K_t^{i,j}$'.

$$K_t^{i,j} = k(O_{t,i}, O_{t,j}) \tag{6}$$

Where,

t - Entities from the feature vector of the training set

Replacing 'K_t' with 'O' in equation (4) leads to optimisation as shown in equation (6).

$$\min_{M_o \in E^{c}\times n_t} \frac{1}{\frac{2}{2}} \|M_o.K_t - O\|^2 + \frac{\varphi}{2} \|M_o\|^2 \tag{7}$$

'M_o' is of dimensions '$c \times n_t$' instead of '$c \times h$' in equation (7), where 'n_t' gives the amount of training feature vectors. This does not affect the solution in equation (6) that may be found using the same replacement employed in equation (5), leading to equation (8).

$$M_o = O(K_t^T K_t + \varphi I)^{-1} K_t^T \tag{8}$$

In EKM, classification is carried out differently in contrast to normal ELM as seen in equation (3). EKM algorithm demands 'M_i' and 'M_o' like ELM, and 'O_t' which is a set of training features in ELM-space. In addition to these parameters, feature classification in 'O' is forwarded to the kernel function (Radial Basis Function (RBF)). A Kernel matrix (K) is built with entries '$K_{i,j} = k.(O_t, i, O_j)$'. The output matrix '$M_o$' is multiplied with the kernel matrix as shown in equation (2), rather than performing classification as shown in equation (3).

$$\hat{O} = M_o.K \tag{9}$$

where,

$\hat{O}$ - Collection of scores assigned to each input feature of every class as shown in equation (3)

Class prediction is performed as shown in equation (4). The proposed system uses EKM for the classification of fish species to attain improved classification rates. The performance of SPTN with EKM is analysed for two publicly accessible datasets and a newly created Cephalopod dataset.

In this work, ELM is implemented using 100 hidden neurons, and performance is compared with EKM variant which employs RBF kernel. The regularization parameter is empirically set to 1e-3, and kernel parameters are tuned via grid search. The primary contribution lies in the application and adaptation of EKM to the domain of marine species classification, particularly Cephalopod identification. To the best of our knowledge, this is the first attempt to apply EKM in this particular context. The methodology is evaluated on two publicly available datasets and a newly created Cephalopod image dataset, demonstrating improved classification accuracy and robustness.

6. Results and discussion

The proposed SPTN-EKM is implemented using Google Colab Pro (Bisong 2019). It is a computational setup that uses a Jupyter notebook with 25 GB of RAM and GPU support. Investigations are done using QUT, Fish-Pak dataset, and a tailored Cephalopod dataset to assess the efficiency and competence of the proposed scheme. To confirm, efficiency is determined using Learning Rate (LR), whereas performance of classification is determined in terms of Precision, Accuracy, Recall and F1-Score. 70% of images are used for training, while 15% are employed for validation, and 15% are used for testing.

The model's accuracy is determined once it is trained using the dataset. Accuracy is taken as the model's significant assessment metric.

$$\text{Accuracy} = \frac{\text{Quanity of rightly categorized fish species}}{\text{Total quantity of fish species}} \tag{10}$$

Loss function which is based on categorical Cross Entropy includes Softmax operations in the last output layer.

$$\text{Loss} = \sum_{i=1}^{n} j = 1\sum^{K} t_{ij} \ln o_{ij} \tag{11}$$

where,

n - Number of fish species

K - Number of fish classes

t_{ij} - i^{th} fish sample

o_{ij} - Output for sample 'i' in class 'j'

Precision, Recall and F1-Score are used for evaluating the performance of the proposed method. In the ensuing section, the inferences of investigations along with the performance of the schemes are discussed.

6.1. Classification of fish species of QUT dataset

Initially, LR is set to a value of 0.01. It drops after every 4^{th} epoch with a decay factor of 0.5. The total number of training epochs is initialised with a value of 100 and 64 batches per epoch are applied. Adam optimizer is used for optimization. Categorical Cross Entropy loss function which integrates Softmax activations in the final output layer, is used for training.

It is necessary to select the optimal SPTN for the QUT dataset. Softmax classifier is used for conducting classification-based investigations for the 6 chosen SPTN frameworks which include Xception, NasNetMobile, DenseNet201, MobileNetV2, InceptionV3 and ResNet50.

Table 3 shows the performance of various classifiers across diverse models on the QUT dataset. Standard classifiers namely, Softmax and ELM are used for confirming the efficacy of the EKM classifier. Xception network is used for extracting features from images. In ELM, the number of hidden nodes is fixed to 512. The algorithms operate within the same dimensional space, thus establishing the dominance of EKM over ELM. Table 4 shows the training time of Xception models on the QUT dataset.

Figure 5 shows the Accuracy offered by the models assessed on the QUT dataset. The performance of the EKM classifier is compared with Softmax and ELM classifiers. It is seen that EKM offers 4.66% and 1.64%, 4.44% and 1.51%, 4.03% and 2.23%, 2.52% and 1.26%, 1.21% and 1.11%, and 2.71% and 1.93% better Accuracy for Xception, NASNetMobile, DenseNet201, MobileNetV2, InceptionV3 and ResNet50 models in contrast to Softmax and ELM classifiers respectively.

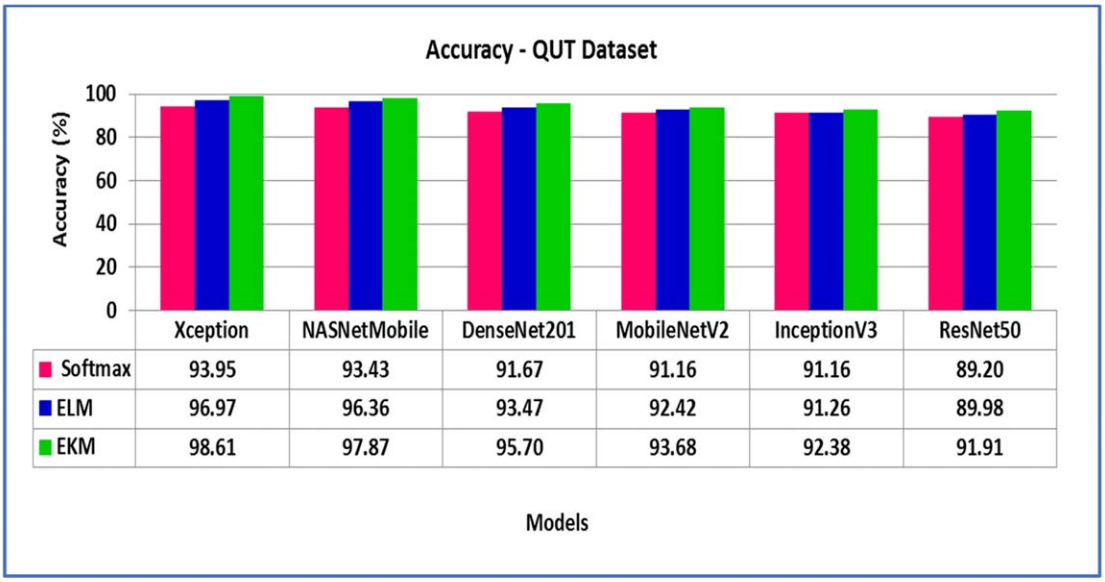


Figure 5. Accuracy of models on the QUT dataset.

Table 3. SPTN for different classifiers on the QUT dataset.

SPTN	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
		Softmax Classifier		
Xception	93.95	97.28	96.08	96.67
NASNetMobile	93.43	96.41	95.88	96.14
DenseNet201	91.67	92.11	95.33	93.69
MobileNetV2	91.16	91.43	94.67	93.02
InceptionV3	91.16	90.74	93.92	92.30
ResNet50	89.20	87.50	92.63	90.46
		ELM Classifier		
Xception	96.97	97.56	97.93	97.74
NASNetMobile	96.36	97.38	97.10	97.24
DenseNet201	93.47	94.73	96.50	95.60
MobileNetV2	92.42	93.58	95.08	94.33
InceptionV3	91.26	91.92	94.96	93.42
ResNet50	89.98	87.97	93.02	90.89
		EKM Classifier		
Xception	98.61	98.90	99.08	98.99
NASNetMobile	97.87	98.54	98.36	98.45
DenseNet201	95.70	95.87	97.71	96.78
MobileNetV2	93.68	94.81	96.46	95.63
InceptionV3	92.38	93.74	95.99	94.85
ResNet50	91.91	90.74	94.29	92.48

Table 4. Training time of Xception models on the QUT dataset.

Classifier/ metric	Training time (Sec)
Softmax	237.62
ELM	25.56
EKM	30.75

Figure 6 shows the Precision offered by the models evaluated on the QUT dataset. It is seen that EKM offers 1.62% and 1.34%, 2.13% and 1.16%, 3.77% and 1.15%, 3.39% and 1.23%, 3% and 1.82%, and 3.24% and 2.77% better Precision for Xception, NASNetMobile, DenseNet201, MobileNetV2, InceptionV3 and ResNet50 models when compared to Softmax and ELM classifiers respectively.

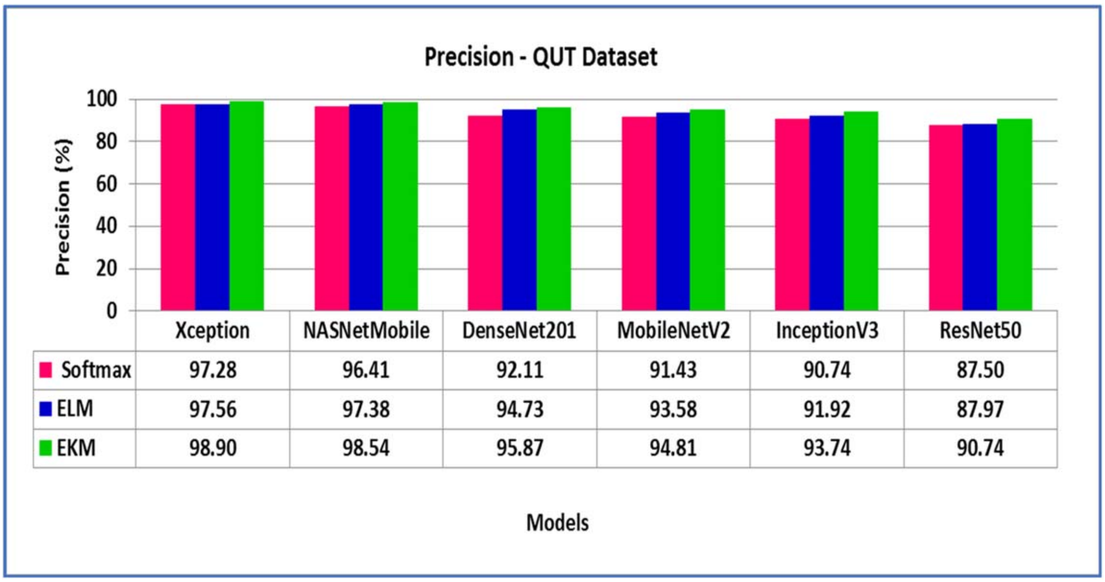


Figure 6. Precision of models on the QUT dataset.

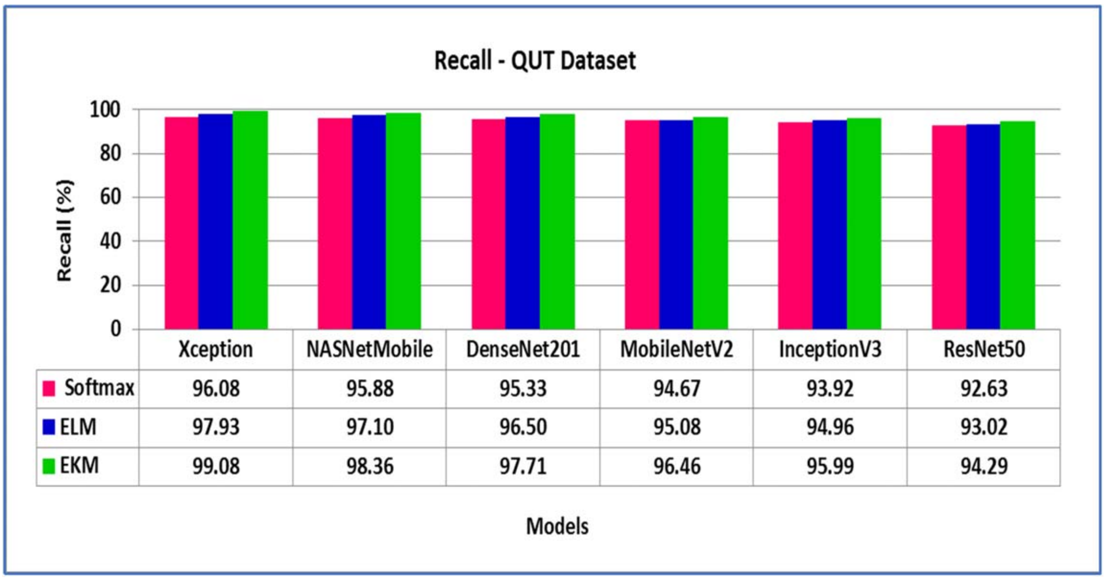


Figure 7. Recall of models on the QUT dataset.

Figure 7 shows the Recall offered by the models assessed on the QUT dataset. It is seen that EKM offers 3% and 1.15%, 2.48% and 1.26%, 2.37% and 1.21%, 1.78% and 1.38%, 1.04% and 1.03%, and 1.66% and 1.27% better Recall for Xception, NASNetMobile, DenseNet201, MobileNetV2, InceptionV3 and ResNet50 models in contrast to Softmax and ELM classifiers respectively.

Figure 8 shows the F1-Score achieved by the models evaluated on the QUT dataset. It is seen that EKM offers 2.32% and 1.25%, 2.30% and 1.21%, 3.09% and 1.18%, 2.61% and 1.29%, 2.55% and 1.43%, and 2.02% and 1.59% better F1-Score for Xception, NASNetMobile, DenseNet201, MobileNetV2, InceptionV3 and ResNet50 models when compared to Softmax and ELM classifiers respectively.

6.2. Classification of fish species of the Fish-Pak dataset

Adam optimizer is employed for optimisation with a batch size of 32 and an initial LR of 0.002. LR decreases every 4 epochs using a decay factor of '0.5'. For experiments carried out using the Fish-Pak dataset, training is conducted for 25 epochs with the number of hidden nodes set to 256. Investigations are carried out to assess the efficacy of SPTN-EKM on the Fish-Pak dataset. The performance of all 6 models is assessed for Softmax, ELM

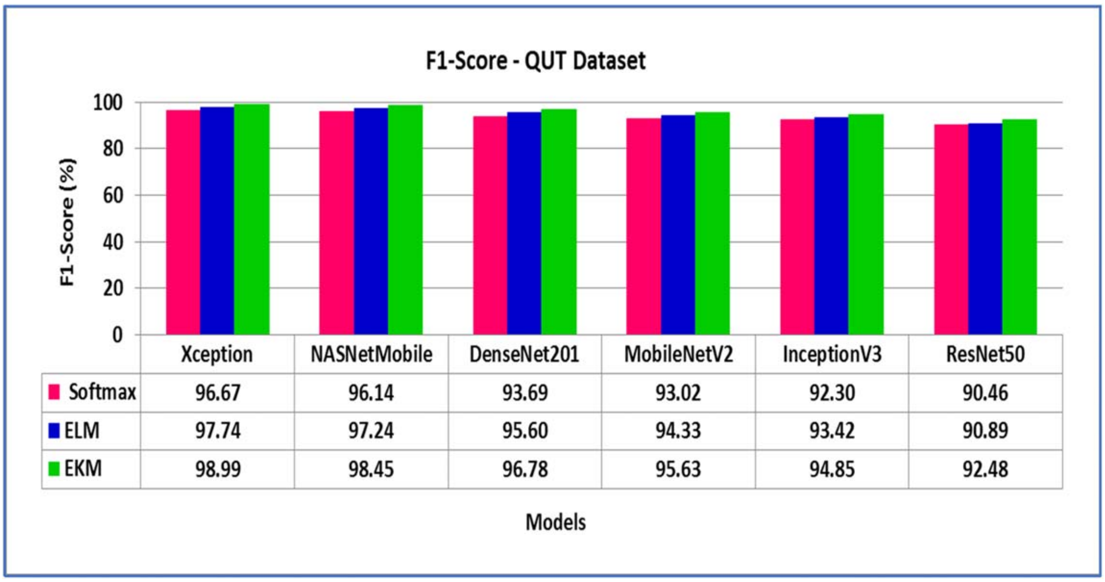


Figure 8. F1-score of models on the QUT dataset.

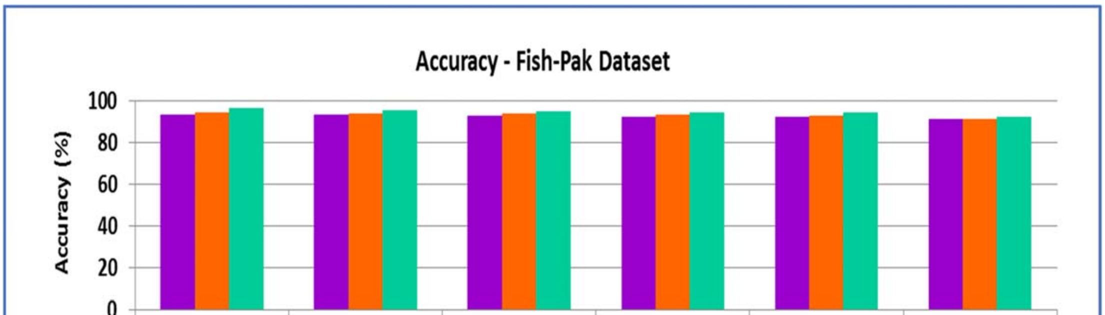
Table 5. SPTN for different classifiers on the Fish-Pak dataset.

SPTN	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
		Softmax Classifier		
Xception	93.13	94.74	96.56	95.64
NASNetMobile	92.92	93.23	96.28	94.73
DenseNet201	92.56	93.18	95.90	94.52
MobileNetV2	92.06	92.91	95.54	94.21
InceptionV3	91.73	92.48	95.33	93.88
ResNet50	90.72	90.06	95.09	92.51
		ELM Classifier		
Xception	94.24	95.45	96.73	96.09
NASNetMobile	93.72	95.02	95.63	95.32
DenseNet201	93.30	94.50	95.26	94.88
MobileNetV2	93.05	94.23	95.24	94.73
InceptionV3	92.37	93.52	94.91	94.21
ResNet50	90.81	91.60	93.94	92.76
		EKM Classifier		
Xception	96.34	97.97	97.86	97.91
NASNetMobile	94.78	96.68	97.58	97.13
DenseNet201	94.70	95.98	97.08	96.53
MobileNetV2	94.09	95.78	97.03	96.40
InceptionV3	93.73	95.09	96.73	95.90
ResNet50	92.06	93.56	96.28	94.90

and EKM classifiers. Table 5 shows the performance of classifiers applied to diverse models for the Fish-Pak dataset. Table 6 provides the training time consumed by DenseNet201Model when applied to the same dataset.

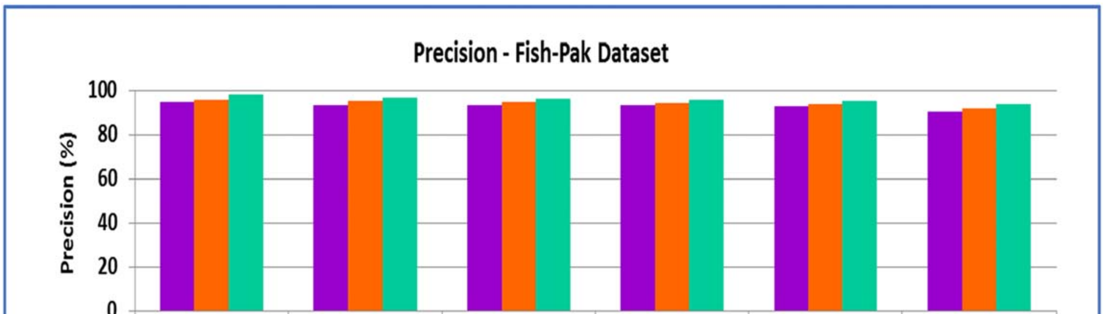
Figure 9 shows the Accuracy achieved by the models evaluated on the Fish-Pak dataset. The performance of the EKM classifier is compared with Softmax and ELM classifiers. It is seen that EKM offers 3.2% and 2.09%, 1.87% and 1.06%, 2.14% and 1.41%, 2.03% and 1.04%, 2% and 1.36%, and 1.34% and 1.25% better Accuracy for DenseNet201, NASNetMobile, MobileNetV2, InceptionV3, Xception and ResNet50 models in contrast to Softmax and ELM classifiers respectively.

Figure 10 presents the Precision offered by the models assessed on the Fish-Pak dataset. It is seen that EKM offers 3.23% and 2.52%, 3.45% and 1.66%, 2.81% and 1.49%, 2.87% and 1.55%, 1.05% and 1.57%, and 1.55% and 1.96% better Precision for DenseNet201, NASNetMobile, MobileNetV2, InceptionV3, Xception and ResNet50 models when compared to Softmax and ELM classifiers respectively.



	DenseNet201	NASNetMobile	MobileNetV2	InceptionV3	Xception	ResNet50
Softmax	93.13	92.92	92.56	92.06	91.73	90.72
ELM	94.24	93.72	93.30	93.05	92.37	90.81
EKM	96.34	94.78	94.70	94.09	93.73	92.06

Figure 9. Accuracy of models on the Fish-Pak dataset.



	DenseNet201	NASNetMobile	MobileNetV2	InceptionV3	Xception	ResNet50
Softmax	94.74	93.23	93.18	92.91	92.48	90.06
ELM	95.45	95.02	94.50	94.23	93.52	91.60
EKM	97.97	96.68	95.98	95.78	95.09	93.56

Figure 10. Precision of models on the Fish-Pak dataset.

Table 6. Training time of DenseNet201model on Fish-Pak dataset.

Classifier/metric	Training time (Sec)
Softmax	185.34
ELM	19.81
EKM	25.23

Figure 11 shows the Recall achieved by the models assessed on the Fish-Pak dataset. It is seen that EKM offers 1.3% and 1.13%, 1.3% and 1.96%, 1.18% and 1.82%, 1.49% and 1.79%, 1.4% and 1.82%, and 1.19% and 2.35% better Recall for DenseNet201, NASNetMobile, MobileNetV2, InceptionV3, Xception and ResNet50 models in contrast to Softmax and ELM classifiers respectively.

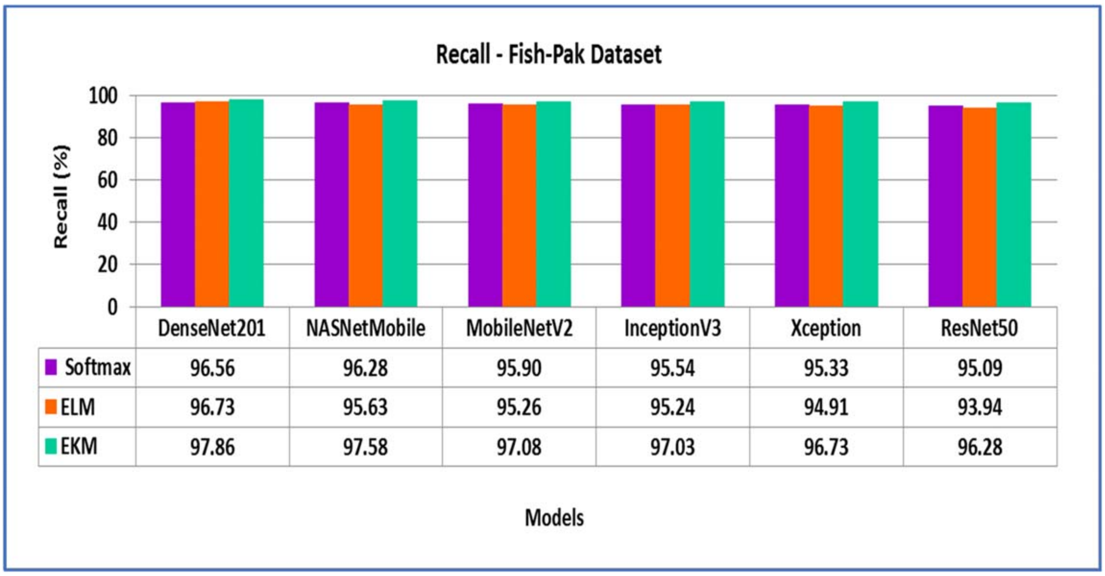


Figure 11. Recall of Models on Fish-Pak dataset.

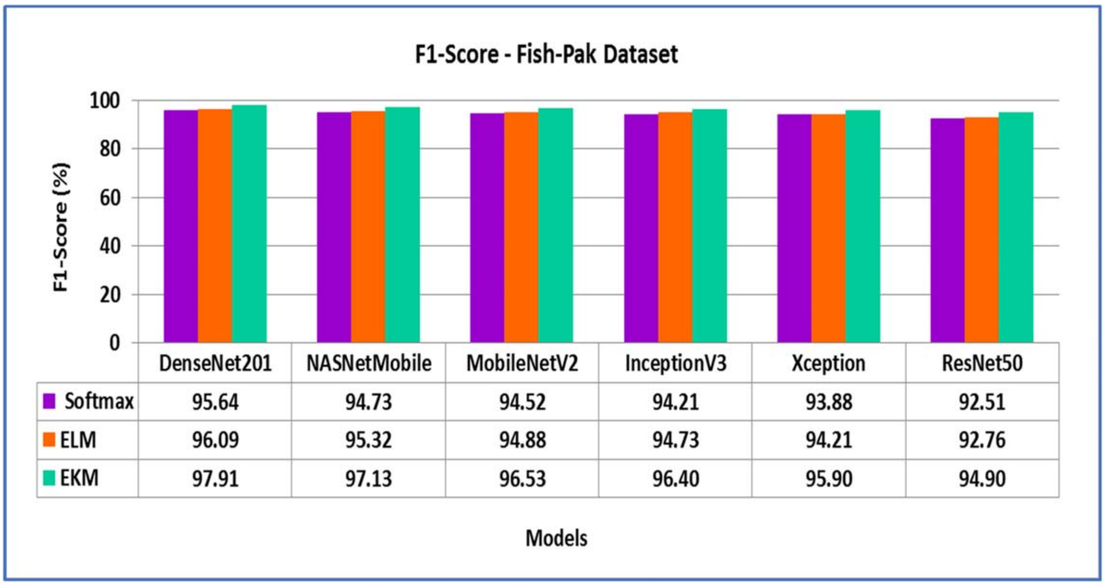


Figure 12. F1-score of models on the Fish-Pak dataset.

Figure 12 shows the F1-Score offered by the models evaluated on the Fish-Pak dataset. It is seen that EKM offers 2.27% and 1.83%, 2.4% and 1.81%, 2.01% and 1.65%, 2.19% and 1.67%, 2.02% and 1.69%, and 2.4% and 2.15% better F1-Score for DenseNet201, NASNetMobile, MobileNetV2, InceptionV3, Xception and ResNet50 models respectively in contrast to Softmax and ELM classifiers.

6.3. Classification of fish species of the Cephalopod dataset

From the results obtained on the Fish-Pak and QUT datasets, it is observed that EKM performs better than other classifiers. SPTNs are applied with EKM for recognising Cephalopods. Table 7 presents the performance of the EKM classifier across diverse models on the Cephalopod dataset. In table 8, it is observed that DenseNet201 achieves a high Accuracy rate of 93.34% with a training time of 27.34 Secs. NASNetMobile is lighter in contrast to DenseNet and offers an Accuracy of 92.3%. MobileNetV2 offers the third highest Accuracy of 91.45%. As NASNetMobile is lightweight and offers high Accuracy, it is well-suited for deployment on IoT-based handheld devices. Though DenseNet201 also yields comparable results, its larger model size may limit its direct use in such environments, unless it is optimized or compressed.

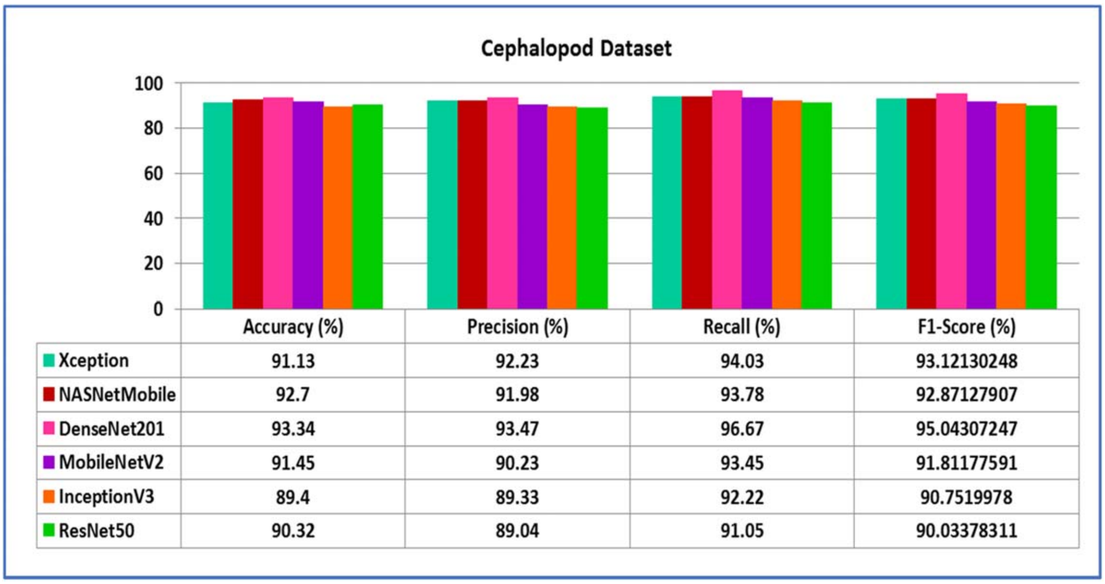


Figure 13. SPTN for different classifiers on the Cephalopod dataset.

Table 7. SPTN for EKM classifier on the Cephalopod dataset.

SPTN	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Xception	91.13	92.23	94.03	93.12
NASNetMobile	92.30	91.98	93.78	92.87
DenseNet201	93.34	93.47	96.67	95.04
MobileNetV2	91.45	90.23	93.45	91.81
InceptionV3	89.40	89.33	92.22	90.75
ResNet50	90.32	89.04	91.05	90.03

Table 8. Training time of DenseNet201model on the Cephalopod dataset.

Classifier/metric	Accuracy (%)	Training time (Sec)
EKM	93.34	27.34

From figure 13, it is evident that DenseNet201 offers 2.21%, 1.04%, 1.89%, 3.94% and 3.02% better Accuracy, 1.24%, 1.49%, 3.24%, 4.14% and 4.43% better Precision, 2.64%, 2.89%, 3.22%, 4.45% and 5.62% better Recall and 1.92%, 2.17%, 3.23%, 4.29% and 5.01% better F1-Score in contrast to Xception, NASNetMobile, MobileNetV2, InceptionV3 and ResNet50 models respectively.

6.4. Overall observations

Based on the outcomes of investigations performed, it is obvious that Xception with EKM is appropriate on the QUT dataset, while DenseNet201 with EKM is suitable for extracting features from the Fish-Pak dataset. ELM and EKM involve reduced training time in contrast to Softmax. EKM consumes some amount of training time as it involves more complex computations in contrast to standard schemes. The proposed scheme supports effective and precise detection of fish species. When combined with a Softmax classifier, Xception outperforms all other models.

Further, ELM offers a significant increase in classification rates. Models with EKM offer further improvement. In case of the Fish-Pak dataset, DenseNet201 when used with Softmax outperforms all other models in terms of classification Accuracy. Use of ELM and EKM offers better results on the QUT dataset. The training time involved by classifiers along with Accuracy is shown in tables 4, 6 and 8. As ELM and EKM are used, the time taken for training is drastically reduced. Table 9 gives the Accuracy and training time of the benchmarked and proposed models involved in classification. Compared to existing models, the proposed model offers increased Accuracy, making it highly suitable for identifying the selected fish species.

Table 9. Classification results of different benchmarked methods.

Methods	Dataset	Accuracy (%)	Training Time
Deep CNN (Iqbal *et al* 2021)	QUT	90.48	—
Improved AlexNet (Ju & Xue 2020)	QUT	97.53	43 min
TL with SE-ResNet152 (Xu *et al* 2021)	Fish-Pak	95.57 (Mean)	—
Proposed Method	QUT	98.61	237 Sec
	Fish-Pak	96.34	185 Sec
	Cephalopod	93.34	27.34 Sec

Training time of the proposed method corresponds to the EKM classifier which is non-iterative and offers significantly faster learning. Softmax and ELM classifiers were used during the initial stages of investigations, but are not included in the final proposed approach.

Table 10. ANOVA test results of the proposed SPTN-EKM scheme with respect to classification accuracy.

Groups	Count	Sum	Mean	Variance		
Proposed SPTN-EKM	30	1748	58.26	21.65		
Xception	30	1562	52.06	22.98		
NASNetMobile	30	1438	47.94	24.64		
DenseNet201	30	1212	40.42	26.79		
MobileNetV2	30	1098	36.64	25.32		
ANOVA						
Source of Variation	Sum of Squares (SS)	degree of freedom (df)	Mean Square (MS)	F	P-Value	F-Critical
Between Groups	5642.86	16	352.68	85.22	3.24532E-22	3.2674
Within Groups	542	131	4.138			
Total	6184.86	150				

Table 11. ANOVA test results of the proposed SPTN-EKM scheme with respect to training time.

Groups	Count	Sum	Mean	Variance		
Proposed SPTN-EKM	30	1978	65.94	21.56		
Xception	30	1764	58.82	22.86		
NASNetMobile	30	1642	54.74	22.18		
DenseNet201	30	1548	51.64	25.68		
MobileNetV2	30	1326	44.22	24.62		
ANOVA						
Source of Variation	Sum of Squares (SS)	degree of freedom (df)	Mean Square (MS)	F	P-Value	F-Critical
Between Groups	5114.86	16	319.68	64.196	2.98764E-21	3.2186
Within Groups	654	131	4.98			
Total	5768.86	150				

6.5. Results of the statistical ANOVA test

Statistical validation of the proposed SPTN-EKM approach and baseline DL models of Xception, NASNetMobile, DenseNet201 and MobileNetV2 are conducted using ANOVA test to assess how effectively the proposed SPTN-EKM approach performs well in terms of classification accuracy and training time. The null and alternate hypothesis used for statistical validation of the proposed SPTN-EKM approach using the ANOVA statistical test is presented as follows.

- **Null Hypothesis** (H_0): There is no significant difference in accuracy and training time following the implementation of the proposed SPTN-EKM approach on Cephalopod classification.
- **Alternate Hypothesis** (H_1): There is a significant difference in accuracy and training time following the implementation of the proposed SPTN-EKM approach on Cephalopod classification.

In specific, tables 10 and 11 present the ANOVA test-based statistical comparisons of classification accuracy and training time of the proposed SPTN-EKM approach for baseline DL models including Xception, NASNetMobile, DenseNet201 and MobileNetV2.

Statistical tests were performed on a validation set with 30 samples. The ANOVA test results of the proposed SPTN-EKM approach depicted in tables 10 and 11 evidently prove that the F-statistic value determined with respect to accuracy and training time is greater than the F-critical value. Specifically, the value of 'p' is identified to be less than 0.05 for both classification Accuracy and training time. Thus, the null hypothesis (H_0) is ignored, and the Alternate Hypothesis (H_1) which asserts that the proposed SPTN-EKM approach offers a significant difference in the classification accuracy and training time on Cephalopod classification is accepted. This confirms that the proposed SPTN-EKM approach performs statistically better than the baseline Xception, NASNetMobile, DenseNet201 and MobileNetV2-based DL models.

7. Conclusion

The proposed SPTN-EKM-based DL mechanism which integrates the strengths of SPTN and EKM offers a high classification rate in recognition of fish species. EKM is confirmed to be an effective classifier, and at the same time, SPTN is identified to support efficient feature extraction. Specifically, incorporation of an optimum SPTN enhances the efficiency of feature learning, enabling automatic and reliable extraction of features with maximized optimality. In contrast, the use of EKM aids in improving the learning efficacy and classification Accuracy. Experiments assessing the performance of the proposed SPTN-EKM model on public QUT and Fish-Pak datasets, and synthetic Cephalopods datasets demonstrate lower computational costs. The results confirm an improved Accuracy of 98.61% on the QUT dataset and better Precision of 96.34% on the Fish-Pak dataset. It also offers an Accuracy of 93.34% on the Cephalopod dataset.

Future scope of enhancement

As a part of the future scope of research, the constructed dataset can be extended to include a wide variety of fish species with environmental conditions to ensure robustness and generalizability of the adopted DL-based model. It is also decided to use the merits of the YOLO Machine Learning (ML) model to accurately detect and classify fish species.

Data availability statement

We would like to clarify that our study makes use of three datasets - two publicly available datasets and one customised dataset developed specifically for our research. The links to the public datasets have been provided in the manuscript for reference. The customised dataset, however, is part of an internal research effort and is currently not prepared for public distribution. We will be happy to share it upon reasonable request, once it is suitably documented and cleared for sharing. The data that support the findings of this study are available upon reasonable request from the authors.

References

Anantharajah K, Ge Z, McCool C, Denman S, Fookes C, Corke P and Sridharan S 2014 Local inter-session variability modelling for object classification *IEEE Winter Conf. on Applications of Computer Vision* 309–16

Ataç D and Şahin D Ö 2024 Evaluation of the performance of transfer learning techniques in classifying fish species *8th Int. Artificial Intelligence and Data Processing Symp. (IDAP)* (IEEE) 1–6

Bisong E 2019 *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (Springer) 59–64

FAO 2012 *Greening the Economy with Agriculture* (FAO)

Fischer J 2014 Fish identification tools for biodiversity and fisheries assessments: review and guidance for decision-makers *FAO Fisheries and Aquaculture Technical Paper* 585 (FAO) I

Gao G, Sun Z, Mu G, Yin H and Ren Y 2024 Research on marine fish classification and recognition based on an optimized ResNet50 model *Marine and Coastal Fisheries* **16** e10317

Huang G B, Wang D H and Lan Y 2011 Extreme learning machines: a survey *International journal of machine learning and cybernetics* **2** 107–22

Iqbal M A, Wang Z, Ali Z A and Riaz S 2021 Automatic fish species classification using deep convolutional neural networks *Wirel. Pers. Commun.* **116** 1043–53

Ishaq O, Sadanandan S K and Wählby C 2017 Deep fish: deep learning–based classification of zebrafish deformation for high-throughput screening *SLAS Discovery: Advancing Life Sciences R&D* **22** 102–7

Jiang L, Quan H, Xie T and Qian J 2022 Fish recognition in complex underwater scenes based on targeted sample transfer learning *Multimedia Tools Appl.* **81** 25303–17

Jin L and Liang H 2017 Deep learning for underwater image recognition in small sample size situations *IEEE OCEANS 2017-Aberdeen* 1–4

Ju Z and Xue Y 2020 Fish species recognition using an improved AlexNet model *Optik* **223** 165499

Karlsson V and Rosvall E 2017 *Student Thesis* KTH, School of Engineering Sciences (SCI)

Li X, Shang M, Qin H and Chen L 2015 Fast accurate fish detection and recognition of underwater images with fast r-cnn *IEEE OCEANS 2015-MTS/IEEE Washington* 1–5

Liu X, Jia Z, Hou X, Fu M, Ma L and Sun Q 2019 Real-time marine animal images classification by embedded system based on mobilenet and transfer learning *IEEE OCEANS 2019-Marseille* 1–5

Liu Z, Cao Y, Li Y, Xiao X, Qiu Q, Yang M and Cui L 2020 Automatic diagnosis of fungal keratitis using data augmentation and image fusion with deep convolutional neural network *Comput. Methods Programs Biomed.* **187** 105019

Mackie I M, Pryde S E, Gonzales-Sotelo C, Medina I, Pérez-Martın R, Quinteiro J and Rehbein H 1999 Challenges in the identification of species of canned fish *Trends in Food Science & Technology* **10** 9–14

Malik H, Naeem A, Hassan S, Ali F, Naqvi R A and Yon D K 2023 Multi-classification deep neural networks for identification of fish species using camera captured images *PLoS One* **18** e0284992

Mathur M, Vasudev D, Sahoo S, Jain D and Goel N 2020 Crosspooled FishNet: transfer learning based fish species classification model *Multimedia Tools Appl.* **79** 31625–43

Montserrat D M, Lin Q, Allebach J and Delp E J 2017 Training object detection and recognition CNN models using data augmentation *Electron. Imaging* **29** 27–36

Park J H, Hwang K B, Park H M and Choi Y K 2019 Application of CNN for fish species classification *Journal of the Korea Institute of Information and Communication Engineering* **23** 39–46

Prabha P A, Suchitra G and Saravanan R 2023 Cephalopods classification using fine tuned lightweight transfer learning models *Intelligent Automation & Soft Computing* **35** 3065–79

Purcell S W, Mercier A, Conand C, Hamel J F, Toral-Granda M V, Lovatelli A and Uthicke S 2013 Sea cucumber fisheries: global analysis of stocks, management measures and drivers of overfishing *Fish and fisheries* **14** 34–59

Qin H, Li X, Liang J, Peng Y and Zhang C 2016 DeepFish: accurate underwater live fish recognition with a deep architecture *Neurocomputing* **187** 49–58

Queensland University of Technology (QCR.CAI) 2016 (https://wiki.qut.edu.au/display/cyphy/Fish+Dataset)

Quinteiro J, Sotelo C G, Rehbein H, Pryde S E, Medina I, Pérez-Martín R I and Mackie I M 1998 Use of mtDNA direct polymerase chain reaction (PCR) sequencing and PCR-restriction fragment length polymorphism methodologies in species identification of canned tuna *J. Agric. Food Chem.* **46** 1662–9

Rauf H T 2019 (https://data.mendeley.com/datasets/n3ydw29sbz/1)

Salman A, Jalal A, Shafait F, Mian A, Shortis M, Seager J and Harvey E 2016 Fish species classification in unconstrained underwater environments based on deep learning *Limnology and Oceanography: Methods* **14** 570–85

Satoto B D, Khotimah B K, Syarief M, Yusuf M, Sophan M K and Anamisa D R 2023 Marine fish species classification using transfer learning and residual network *2023 IEEE 9th Information Technology Int. Seminar (ITIS)* 1–6

Shah S Z H, Rauf H T, IkramUllah M, Khalid M S, Farooq M, Fatima M and Bukhari S A C 2019 Fish-Pak: fish species dataset from Pakistan for visual features based classification *Data in brief* **27** 104565

Shang H, Yu Y and Song W 2023 Underwater fish image classification algorithm based on improved ResNet-RS model *Jiangsu Annual Conference on Automation (JACA 2023)* 106–10

Siddiqui S A, Salman A, Malik M I, Shafait F, Mian A, Shortis M R and Harvey E S 2018 Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data *ICES J. Mar. Sci.* **75** 374–89

Sirigineedi M, Jagan Mohan R N V and Sahu B 2024 Improving fish image detection speed with hybrid VGG16 and darknet *Multimedia Tools Appl.* **84** 1–16

Thilsted S H, Thorne-Lyman A, Webb P, Bogard J R, Subasinghe R, Phillips M J and Allison E H 2016 Sustaining healthy diets: the role of capture fisheries and aquaculture for improving nutrition in the post-2015 era *Food Policy* **61** 126–31

Winoto A S, Kristianus M and Premachandra C 2020 Small and slim deep convolutional neural network for mobile device *IEEE Access* **8** 125210–22

Xu X, Li W and Duan Q 2021 Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification *Comput. Electron. Agric.* **180** 105878

Zaneveld J R, Burkepile D E, Shantz A A, Pritchard C E, McMinds R, Payet J P and Thurber R V 2016 Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales *Nat. Commun.* **7** 1–12

Zhang S, Liu W, Zhu Y, Han W, Huang Y and Li J 2022 Research on fish identification in tropical waters under unconstrained environment based on transfer learning *Earth Sci. Inf.* **15** 1155–66