RESEARCH ARTICLE

Design and Implementation of a Software Engineering-Driven Deep Transfer Learning Framework for Seafood Fish Detection

Muzzamal Hameed[1], **Muhammad Haroon**[2], **Amna Farooq**[3,*] and **Abdul Karim Sajid Ali**[4]

[1] Department of Computer Science, School of Electrical Engineering & Computer Science, National University of Sciences and Technology (NUST), Islamabad, Pakistan
[2] Department of Computer Science, Riphah School of Computing & Innovation, Riphah International University, Lahore, Pakistan
[3] Department of Computer Science, Government College University, Faisalabad, Sahiwal Campus, Sahiwal 57000, Pakistan
[4] Department of Computer Science, Illinois Institute of Technology, Chicago, IL, 60616, United States

Abstract

Seafood quality inspection is critical for ensuring food safety and minimizing economic losses from spoilage. While traditional methods are slow and labor-intensive, computer vision and machine learning have emerged as efficient automated alternatives. This study presents SFFDNet, a software engineering-driven convolutional neural network featuring a lightweight 19-layer architecture with optimized feature extraction blocks and regularization strategies. With only 2.49 million parameters—significantly fewer than VGG16 (138M) and ResNet50 (25.6M)—our model achieves 98.80% accuracy on the Large-Scale Fish Segmentation and Classification Dataset. SFFDNet outperforms both transfer learning models (VGG16: 96.54%, ResNet50: 53.34%, InceptionV3: 58.39%) and conventional approaches (CNN: 96.00%, SegNet: 88.69%, YOLO+ResNet50: 91.64%). The framework emphasizes computational efficiency, modularity, and scalability, bridging high-accuracy deep learning with practical industrial seafood inspection through software engineering principles.

Keywords: feature extraction, image segmentation, task-specific analysis, quality evaluation, color-based image analysis, classification, food quality inspection, convolutional neural networks, SFFDNet.

1 Introduction

Seafood quality inspection is critical for ensuring food safety and preventing economic losses caused by spoilage. Signs of deterioration and decomposition vary across species, making species-specific inspection essential. However, industrial seafood packaging and processing still rely heavily on manual grading and visual assessment, which are inconsistent, labor-intensive, and often achieve low accuracy. These limitations create a pressing need for automated,

Submitted: 25 July 2025
Accepted: 27 September 2025
Published: 03 November 2025

Vol. 1, **No.** 2, 2025.
10.62762/JSE.2025.535801

***Corresponding author:**
Amna Farooq
amnaf4916@gmail.com

Citation
Hameed, M., Haroon, M., Farooq, A., & Ali, A. K. S. (2025). Design and Implementation of a Software Engineering-Driven Deep Transfer Learning Framework for Seafood Fish Detection. *ICCK Journal of Software Engineering*, 1(2), 109–123.

scalable, and reliable inspection frameworks tailored to seafood quality evaluation [1].

Fish detection is the identification and classification of fish species using shape, texture, and color recognition. It is important in several fields, including, but not limited to, marine biology, ecological monitoring, and the fishing industry. Timely and accurate species recognition is critical for identification and enables informed decisions for trade, conservation, and harvesting. As summarized in, computer vision approaches have been successful in analyzing and recognizing fish by vision using species comparisons. As the consumption of fish increases due to people changing to fish and seafood out of a desire to eat healthier and to avoid red meat, the need for an accurate assessment of its quality has risen. As highlighted by recent FAO reports, approximately 35% of global fish and seafood production is lost or wasted annually, resulting in substantial economic losses amounting to billions of dollars and posing significant public health risks through potential foodborne outbreaks [2]. These statistics emphasize the critical importance of developing automated and reliable seafood inspection frameworks that can reduce waste, ensure food safety, and support sustainable practices within the seafood industry. Unfortunately, in seafood processing, quality inspection continues to rely heavily on manual grading and visual assessment. However, such approaches are often inconsistent, labor intensive, and highly prone to human error, with reported accuracy rates frequently falling below 70% [3]. Furthermore, manual inspection inherently limits throughput, making it unsuitable for meeting the demands of large-scale industrial operations. These shortcomings highlight a critical gap in current practices and underscore the pressing need for automated, scalable, and accurate inspection frameworks that can ensure both efficiency and reliability in seafood quality assurance.

The most recent advances in image processing, pattern recognition, and related areas of research have had an impact on this field. CNNs or Convolutional Neural Networks, for example, have been very effective in image classification problems. Their effectiveness is further augmented by methods like transfer learning, which involve the usage of pre-trained generic models in specialized applications in which the underlying concepts are transferable [4].

Existing fish detection and classification models face two major limitations when applied to seafood quality inspection. First, many architectures such as VGG16 (138M parameters) and ResNet50 (25.6M parameters) are computationally heavy, making them unsuitable for real-time deployment in resource-constrained industrial environments. Second, most are generic image classifiers, not tailored to seafood-specific features such as subtle texture or color changes that indicate spoilage. These shortcomings highlight the need for a lightweight, domain-adapted architecture that balances efficiency and accuracy, motivating the development of SFFDNet.

Despite these advances, a significant research gap remains: most existing CNN models are either excessively large, demanding substantial memory and computational power, or are generic image classifiers that fail to capture seafood-specific visual features such as subtle texture changes indicating morbidity or decomposition. For example, VGG16 (138M parameters) and ResNet50 (25.6M parameters), though highly successful in general-purpose image recognition, require substantial GPU memory and incur inference times unsuitable for high-throughput inspection lines. Their reliance on deep architectures with millions of parameters leads to higher latency and energy consumption, making them impractical for real-time seafood monitoring where decisions must be made within milliseconds. Moreover, these models were not originally designed to detect seafood-specific cues such as fine-grained texture degradation, which are critical for accurate quality assessment. As a result, such architectures are poorly suited for resource-constrained industrial environments where accuracy and efficiency must be jointly optimized. Addressing this challenge requires a domain-adapted, lightweight model that minimizes computational overhead while preserving detection accuracy motivating the development of a novel Software Engineering Driven Fish Detection Network (SFFDNet) based on Convolutional Neural Network (CNN).

To address this gap, we conducted two complementary investigations: (1) a benchmarking study using ten widely adopted pre-trained CNN models to evaluate their performance on seafood fish classification, and (2) the design of a novel lightweight CNN, SFFDNet, specifically optimized for seafood inspection. Unlike the transfer learning models, which rely on large generic architectures, SFFDNet employs a compact 19-layer design with only 2.49M parameters, enabling real-time performance while preserving high accuracy.

Prior studies have often relied on very deep architectures, spanning 40–100+ layers and requiring tens of millions of trainable parameters. While these models achieve strong accuracy, they come at the expense of high computational demand and significant inference latency, which restricts their usability in real-time industrial settings. In contrast, SFFDNet is designed as a lightweight 19-layer network with only 2.49 million parameters, striking a balance between efficiency and accuracy. This compact design enables lower computational cost, faster inference, and reduced latency, making SFFDNet highly practical for real-time seafood inspection pipelines in industrial environments. The architecture of this model is tailored toward seafood species to facilitate their recognition in images, seeking to achieve optimal classification efficiency. To ensure uniformity for both the transfer learning models and the SFFDNet model, all images undergo the same preprocessing sequence, which includes uniform resizing to 255 x 255 pixels and normalization. These techniques are evaluated using a publicly available, large-scale dataset of fish images. From a software engineering perspective, the models are developed with attention to modular design, reusability, and scalability. This allows for easy integration into real-time systems used in seafood quality inspection, contributing to reproducibility and system maintainability.

The remainder of this paper is structured as follows: Section II reviews related work in fish detection using machine learning techniques. Section III details the methodology, including model architecture and training procedures. Section IV presents the dataset, experimental setup, and evaluation results. Section V concludes the study and outlines potential directions for future research. To summarize, the key contributions of this paper are as follows:

1. We propose SFFDNet, a compact CNN specifically tailored for seafood fish detection and classification. With only 2.49M parameters, it achieves higher accuracy than heavy architectures such as VGG16 (138M) and ResNet50 (25.6M).
2. Guided by software engineering principles of modularity, reproducibility, and scalability, SFFDNet is designed for real-time deployment in resource-constrained environments, with demonstrated efficiency in training, inference, and memory usage.
3. SFFDNet was rigorously benchmarked against ten state-of-the-art CNN models, consistently achieving superior accuracy while requiring significantly fewer computational resources.

2 Related Work

The automation of fish quality and species analysis has progressed from early traceability and machine vision systems to modern deep learning. We therefore group related work into two themes: (1) traditional methods for quality assessment and (2) deep learning for detection and classification, providing a clear view of the field's evolution.

2.1 Traditional methods for quality assessment

The use of technology for assessing quality and enabling traceability for seafood and other food products has undergone research from several scholars. Wang et al. [5] developed a food traceability system that employs artificial neural network (ANN) methodologies to ascertain food items' quality grade, showcasing the automation possibilities food evaluation algorithms may offer.

Tao et al. [6] stressed the importance of vision systems in the seafood industry for measurement of visual quality attributes because of the need for objectivity and consistency. Their research demonstrates that computer vision, when coupled with predictive algorithms, may serve in the assessment of appearance-related attributes such as size, shape, color, and texture. These systems could also be enhanced for odor, bone, defect, and irregularity detection.

Linsen et al. [7] created a complete quality traceability and information management system for the beef industry, which was implemented using object-oriented programming. Their study serves as an example of the application of software engineering techniques for the improvement of food safety and food safety monitoring. In a subsequent study, Wang et al. [8] developed a seafood traceability system with built-in anti-counterfeiting features to safeguard the integrity of the seafood and ensure transparency from the point of capture to the end user. Yan et al. [9] created an IoT-based real-time monitoring system for aquaculture ecosystems, enhancing operational agility and data-based decision making in fish farming. In the same way, Wang et al. [10] applied a seafood quality traceability system using the bill of lots method, which enhances seafood traceability and accountability in the distribution network.

Nicolae et al. [11] designed a user-oriented traceability

Table 1. Summary of existing seafood and fish classification studies.

Study	Dataset	Technique	Performance	Computational Cost	Shortcomings
[15]	A Large-Scale Fish Dataset	Shallow deep learning model	94%	Low–Moderate (fewer layers, relatively lightweight)	Limited to 9 species; tested only on public dataset; real-world deployment not validated.
[16]	FFE dataset	Deep learning (SqueezeNet, VGG19) for feature extraction + ML classifiers (k-NN, RF, SVM, LR, ANN)	VGG19 + ANN achieved 77.3%	High (VGG19 feature extraction is computationally expensive)	Accuracy is relatively low; only eye-based features are used; the dataset is limited.
[17]	Mackerel, Sardine, Prawn, Pomfret, Red Snapper, Cuttlefish	Paper-based pH sensor (Methyl Red & Bromocresol Purple) + Random Forest model trained on protein, lipid, and TVB-N values	Pomfret (MSE = 0.004625, RMSE = 0.068007, MAE = 0.065833) and Mackerel (MSE = 0.005034, RMSE = 0.070949, MAE = 0.062933)	Low (sensor + RF is computationally efficient)	Limited to six seafood types; model performance may vary with different storage/packaging; not generalized to all seafood.
[20]	A Large-Scale Dataset for Fish Segmentation and Classification	Deep learning-based image segmentation	Achieved 88.69% segmentation accuracy	High (SegNet has a large encoder–decoder structure)	Focused only on segmentation, not a full classification pipeline; may need more robust testing on diverse underwater conditions
[21]	F4K dataset	Machine learning-based classification	Achieved 93.80% classification accuracy	Moderate (ML-based, less costly than CNNs)	Limited to a specific dataset; may require validation on larger, more diverse datasets for generalization.
[22]	AQUARIO28E40I dataset	CNN	Achieved 96% classification accuracy	Moderate–High (CNN with deeper layers increases cost)	Dataset size limited; may face generalization issues in real-world underwater environments
[23]	LifeCLEF 2015 (Fish4Knowledge)	YOLO with ResNet50	Achieved 91.64% classification accuracy	High (YOLO + ResNet50 is heavy, high GPU requirement)	Performance drops with other datasets (e.g., the UWA dataset achieved only 79.80%)
[24]	Self-collected dataset	Random Forest (RF) with feature points	Achieved 87.30% classification accuracy	Low (tree-based methods are efficient)	Feature engineering-dependent; lacks robustness across diverse species
[25]	Global Information System (GIS) on Fishes	BP (Backpropagation) algorithm	Achieved 82.25% classification accuracy	Low–Moderate (depends on hidden layers, but lighter than CNNs)	Lower accuracy compared to deep learning models; sensitive to noisy data

system for stakeholders in the fish and seafood supply chain. Their focus was on system functionality and adaptability to ensure ease of use across various operational roles. Lastly, Cui et al. [12] developed a method for fish identification in complex underwater terrains, illustrating the application of image-based classification in low visibility conditions like cloudy or murky waters.

2.2 Deep learning for detection and classification

Deep learning has significantly advanced the field of computer vision by enabling models to automatically learn abstract feature hierarchies directly from raw image data [13]. This has reduced dependence on manual feature engineering and improved generalizability across domains. Its effectiveness has been widely demonstrated not only in image classification tasks but also in applications such as object detection [14] and medical diagnosis. These advancements validate the cross-domain power of deep neural networks and motivate their tailored application in areas like seafood quality inspection and fish classification.

Despite notable progress in fish classification and traceability systems through the use of machine learning and image-based methods, relatively fewer studies have directly addressed the challenges of seafood quality inspection using deep learning. For example, Simonyan et al. [18] explored eye-based freshness detection, while Tan et al. [19] integrated biochemical sensors with machine learning to assess spoilage in selected species. These works, while important, were constrained by dataset size, limited feature scope, or reliance on auxiliary sensors. Other studies, such as Kumaravel [17], proposed deep learning–based seafood monitoring but did not explicitly address deployment in real-time industrial pipelines. A summary of existing seafood and fish classification studies, highlighting datasets, techniques, performance, computational cost, and key shortcomings, is shown in Table 1.

The proposed SFFDNet addresses these shortcomings by combining a compact 19-layer architecture (2.49M parameters) with optimized preprocessing. Resizing and normalization enhance robustness under low-light conditions, while hierarchical convolutional blocks improve sensitivity to fine-grained texture differences that signal degradation, even when partial occlusion is present. By balancing accuracy, efficiency, and robustness to visual variability, SFFDNet directly tackles the gaps identified in previous literature, offering a domain-adapted framework for scalable seafood inspection.

From the studies reviewed and summarized in Table 1, several trends are evident. Traditional methods demonstrate the feasibility of automation but remain limited in scalability and robustness. Deep learning approaches achieve higher accuracy but often rely on very large models (e.g., VGG, ResNet, YOLO) with millions of parameters, resulting in high computational cost and limited real-time applicability. Moreover, most existing models are designed for generic fish identification or segmentation rather than seafood-specific quality inspection, and few explicitly address the requirements of industrial integration, such as modularity, reproducibility, and low memory usage. Collectively, these gaps highlight the need for a lightweight yet accurate CNN architecture that is domain-adapted for seafood inspection and designed with software engineering principles to ensure maintainability and deployment readiness. These insights directly motivate the development of our proposed model, SFFDNet.

To bridge this gap, our study introduces a novel deep learning model based on CNN and named SFFDNet, which is specifically tailored for seafood fish classification. It not only outperforms well-established pre-trained models but also offers a significantly lower computational footprint, making it ideal for real-time, embedded, or resource-constrained industrial applications. This work contributes a unique synergy between accuracy, efficiency, and engineering scalability dimensions rarely addressed simultaneously in the existing literature.

3 Methodology

To enable effective detection of various seafood fish species, the input images are resized to 255 $\times$ 255 pixels, ensuring uniformity and compatibility with the model architecture. This resolution was selected after comparative trials with smaller (75 $\times$ 75, 128 $\times$ 128) and larger (512 $\times$ 512) dimensions, balancing computational efficiency with preservation of fine-grained texture details. This helps to standardize the data and reduce processing time. Then, ResNet preprocessing is applied to the resized images. After preprocessing, the images are passed to CNN models for classification. These models are trained to identify and classify the fish into nine distinct categories. The CNN extracts important features from the images and uses them to detect the correct fish type. The complete Workflow diagram illustrating the seafood fish detection and classification process based on the SFFDNet model is shown in Figure 1.

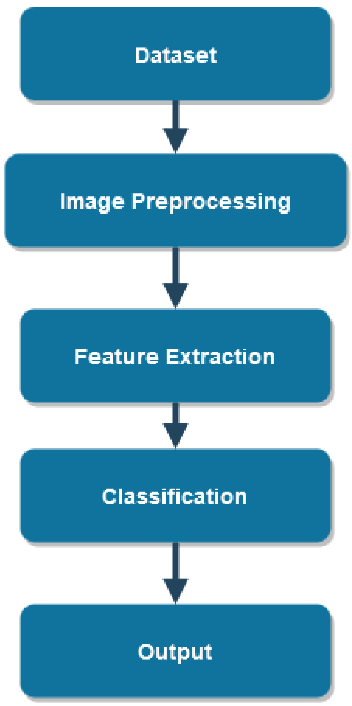


Figure 1. Workflow diagram of the SFFDNet-based seafood fish detection and classification pipeline, explicitly highlighting the major stages: image acquisition, preprocessing (resizing and normalization), feature extraction using the SFFDNet convolutional blocks, and final classification into nine seafood species.

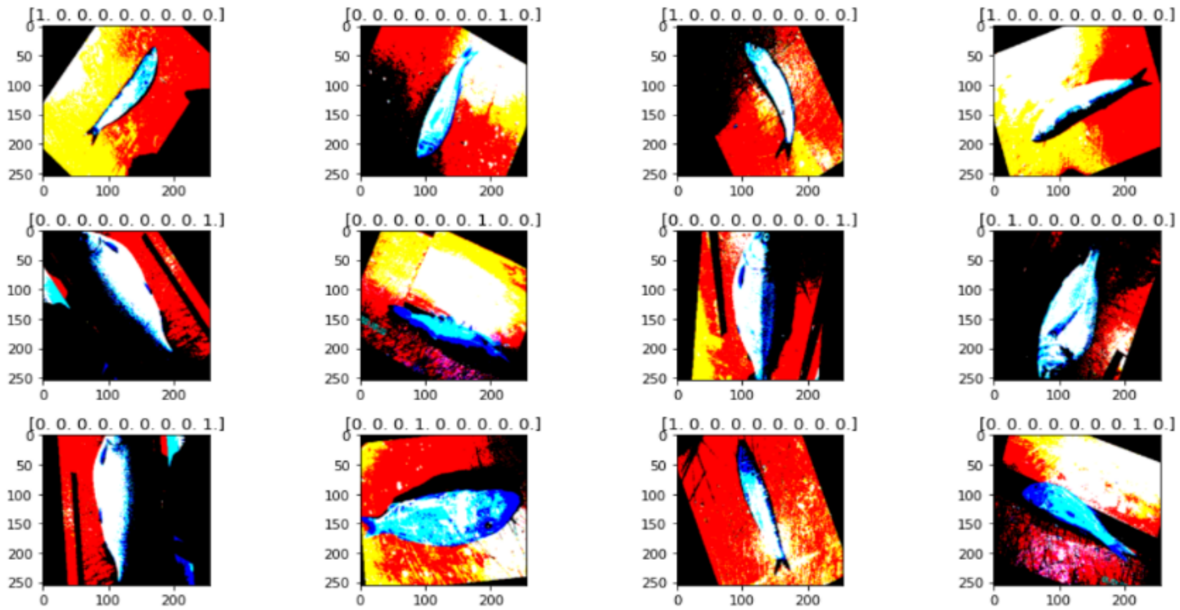


Figure 2. Examples of seafood fish images following the preprocessing stage.

3.1 Dataset preprocessing

The dataset utilized in this study comprises RGB images of seafood fish, originally available in various resolutions and dimensions. To ensure uniformity and optimize computational efficiency, all images were resized to 255×255 pixels before being input into the CNN models. This preprocessing step also aids in stabilizing the training process and improving model performance. A selection of preprocessed sample images from the dataset is presented in Figure 2 to illustrate the standardized input format. The dataset was divided into three subsets: 70% for training, 20% for validation, and 10% for independent testing. The validation set was used to tune hyperparameters and monitor training performance, while the test set was kept completely separate and used only for final evaluation and reporting of accuracy. To standardize the dataset before training, all seafood fish images were normalized and resized, and representative examples after preprocessing are shown in Figure 2.

3.2 Convolutional Neural Network

CNN is a deep learning-based approach that has led to significant advancements in computer vision. In the proposed method, we develop a new CNN model named SFFDNet, which consists of 19 layers and has 2,490,697 total parameters. GlobalAveragePooling2D is applied after the final convolutional block, followed by a Dense(5457) layer, then a Dense(128) layer, and finally the output Softmax layer with 9 units. This design ensures effective feature extraction and accurate classification. The SFFDNet model is trained using various parameter settings, as shown in Table 3. The most suitable parameters are highlighted in bold. The overall architecture of the proposed SFFDNet model is presented in Figure 3, highlighting its hierarchical convolutional blocks, fully connected layers, and output stage for multi-class classification.

Following the convolutional and pooling stages, a Global Average Pooling (GAP) layer was applied to reduce the spatial dimensions. This operation produced a flattened feature vector of length 5,457, which directly informed the size of the subsequent Dense(5457) layer. While the number may appear unconventional compared to round-layer dimensions (e.g., 512, 1024), it is a natural consequence of the feature map dimensions output by the final convolutional block. Retaining this dimensionality allowed the model to preserve all extracted features before classification, thereby maintaining high accuracy without additional compression.

The ten pre-trained models (VGG16, VGG19, EfficientNetB7, DenseNet201, MobileNetV2, Xception, InceptionV3, ResNet50/101/152) were selected due to their widespread adoption and proven effectiveness in prior image classification and fish detection tasks. They represent a diverse range of architectures (from lightweight MobileNet to deeper ResNets and EfficientNet), allowing for a comprehensive comparison of our proposed model against both lightweight and computationally intensive state-of-the-art approaches. We also train 10 CNN models, such as VGG16 [26], VGG19 [26],

Table 2. Layer-by-layer architecture of the proposed SFFDNet model.

Layer Type	Kernel Size / Stride	No. of Filters	Output Shape (H×W×C)	Notes
Input & Rescaling	–	–	255 × 255 × 3	RGB image input
Conv Block 1	3×3 / 1	32	255 × 255 × 32	Conv2D + ReLU + MaxPool
Conv Block 2	3×3 / 1	64	127 × 127 × 64	Conv2D + ReLU + MaxPool
Conv Block 3	3×3 / 1	128	63 × 63 × 128	Conv2D + ReLU + MaxPool
Conv Block 4	3×3 / 1	256	31 × 31 × 256	Conv2D + ReLU + MaxPool
GlobalAveragePooling2D	–	–	256	Replaces Flatten to reduce parameters
Dense Layer 1	–	–	5457	Fully connected, ReLU activation
Dense Layer 2	–	–	128	ReLU activation
Dropout	–	–	128	p = 0.5
Output Layer	–	–	9	Softmax classifier

EfficientNetB7 [27], DenseNet201 [28], MobileNetV2 [29], Xception [30], InceptionV3 [31], Resnet50 , Resnet101, and Resnet152 [32], to compare our model performance with the pre-trained models. The layers of the pre-trained models were modified based on the requirements of our dataset. Specifically, the input layers were adjusted to match the resized image dimensions (255 × 255), ensuring compatibility with the input data. The final layers were replaced to accommodate the nine output classes in the dataset. Table 2 details the full architecture of SFFDNet, showing kernel sizes, filter counts, strides, and output dimensions for each layer. This design emphasizes compactness (2.49M parameters) while retaining strong feature extraction capacity through four optimized convolutional blocks. Together, these architectural refinements enable SFFDNet to balance accuracy with computational efficiency, supporting reproducibility and deployment in resource-constrained environments.

Figure 4 shows the end-to-end workflow of SFFDNet, beginning with preprocessing, followed by CNN feature extraction and final classification. While previously described as a 'two-stage' process, this is more accurately a single unified pipeline consisting of training and inference steps.

Highlighted key aspects of the proposed framework are as follows:

- Training (ResNet model, image preprocessing, feature extraction).
- Classification (unseen images, inference with the model, nine species classification as output).

For clarity and reproducibility, the training and inference process of SFFDNet is summarized below. This workflow demonstrates the preprocessing, optimization, and evaluation pipeline used in this study.

Input: Seafood fish dataset D (resized to 255×255)
Output: Trained SFFDNet model for 9-class classification

1. **Preprocessing:**
   - Resize to 255×255, normalize pixels
   - Apply augmentation(rotation, flip, brightness)
2. **Model (SFFDNet):**
   - Input: 255×255×3
   - 4 Conv blocks: filters [32, 64, 128, 256], kernel 3×3, ReLU, MaxPool
   - GlobalAveragePooling2D
   - Dense(5457, ReLU) → Dropout(0.5) → Dense(128,ReLU)
   - Output: Dense(9, Softmax)
   - Optimizer: Adam (lr=0.001), Loss: Sparse Categorical Cross-Entropy
3. **Training:**
   - Epochs= 20, Batch size = 32
   - Forward → Loss → Backprop → Weight update
   - Early stopping on validation loss
4. **Evaluation:**

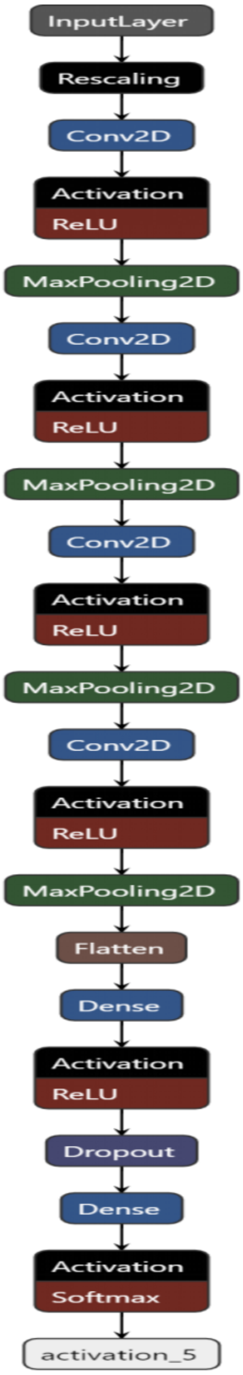


Figure 3. Architectural overview of the proposed SFFDNet model illustrating its 19-layer hierarchical structure, including convolutional blocks, pooling layers, global average pooling, fully connected layers, and the final softmax classifier.

- Metrics: Accuracy, Precision, Recall, F1-score
- Confusion Matrix for class-wise analysis

5. **Inference:**
   - Preprocess image → Forward pass → Predicted species

Each model was trained using various parameter settings, including changes in learning rate, batch size, and optimizer. Hyperparameter selection was carried out using a grid search approach across combinations of learning rates, batch sizes, and optimizers. A validation set (20% of training data) was used to evaluate performance under each configuration. The optimal parameters (batch size = 32, learning rate = 0.001, Adam optimizer, sparse categorical cross-entropy) were chosen based on the highest validation accuracy and lowest validation loss, ensuring generalizability. These configurations are summarized in Table 3, where the most effective settings are highlighted in bold. These optimal settings were selected based on the performance achieved during training.

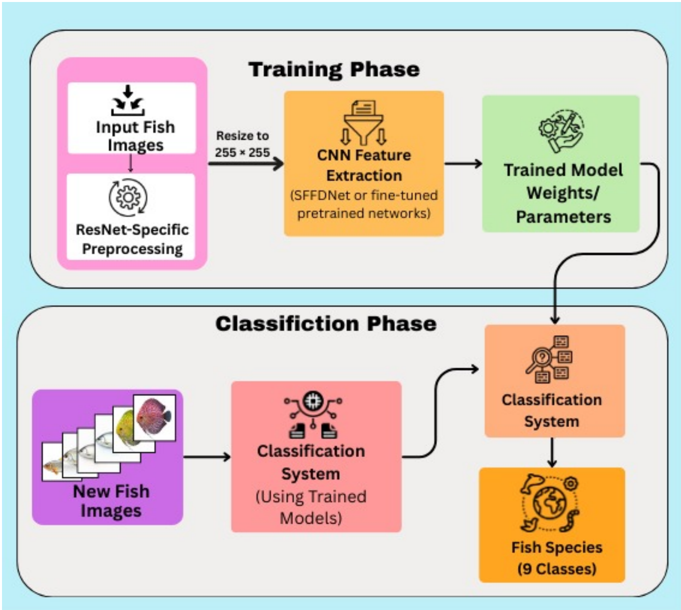


Figure 4. Workflow of the seafood fish classification process, illustrating preprocessing, CNN-based feature extraction, and classification into nine species.

Hyperparameter tuning was performed using a grid search strategy to systematically evaluate multiple configurations. Batch sizes of 32, 64, and 128 were tested, along with learning rates of 0.0001, 0.001, and 0.01, and three optimizers: SGD, RMSProp, and Adam. The evaluation was carried out using a validation split of 20% from the training data. Results showed that smaller batch sizes (e.g., 32) provided stable convergence but required longer training time, while larger sizes (e.g., 128) reduced generalization and increased memory usage. A batch size of 32 was therefore selected as a balanced choice for accuracy and efficiency in our setup. Similarly, a learning rate of 0.001 yielded the most reliable convergence; lower values slowed training, whereas higher values led to unstable loss. Among optimizers, Adam consistently provided the fastest convergence and lowest validation loss compared with SGD and RMSProp. The final hyperparameters, shown in Table 3, represent the configuration that achieved the best trade-off between accuracy, stability, and computational cost.

Table 3. Hyperparameter settings used for training SFFDNet and the baseline CNN models.

Parameters	Settings
Mini Batch Size	32
Number of Epochs	20
Optimizer	Adam
Loss function	Sparse Categorical Crossentropy (SCC)
Learning Rate	0.001

4 Experiments

4.1 Dataset

In this study, we utilize the publicly available Fish dataset titled "A Large-Scale Dataset for Fish Segmentation and Classification", introduced by Oğuzhan Ulucan [20]. The dataset comprises a total of 9,000 images, evenly distributed across nine fish species. Each class contains 1,000 images, ensuring a balanced dataset for training and evaluation. The species included in the dataset are: Black Sea Sprat, Gilt-Head Bream, Horse Mackerel, Red Mullet, Red Sea Bream, Sea Bass, Shrimp, Striped Red Mullet, and Trout. A selection of representative images from each class is illustrated in Figure 5.

The preprocessing pipeline also played a central role in model efficiency and accuracy. Images were resized to 255 $\times$ 255 pixels after comparative trials with smaller and larger dimensions. This resolution was intentionally selected as it preserved fine-grained features such as texture and color cues, which are critical for distinguishing morphologically similar species, while avoiding the high computational burden of larger inputs. Smaller resolutions (e.g., 75 $\times$ 75 or 128 $\times$ 128) accelerated training but consistently led to poorer recognition of subtle spoilage patterns, whereas 512 $\times$ 512 inputs significantly increased memory usage and training time without corresponding accuracy gains. Thus, the choice of 255 $\times$ 255 offered an effective compromise that supports both real-time deployment and accurate feature extraction, aligning with the practical requirements of seafood inspection pipelines.

4.2 Performance metrics

To assess the effectiveness and performance of the proposed system, the evaluation metrics Accuracy (Acc) and Loss are utilized. The equation used to calculate Accuracy is mentioned below

$$\text{Acc} = \frac{(\alpha + \beta)}{(\alpha + \beta + \gamma + \eta)} \tag{1}$$

where α and β represent the number of true positives (TP) and true negatives (TN), respectively, indicating correctly classified positive and negative samples. Similarly, γ denotes the number of false positives (FP), and η denotes the number of false negatives (FN), representing negative and positive samples that were incorrectly classified.

5 Result

In our proposed method, we implemented an advanced CNN-based model, SFFDNet, and 10 pre-trained models to detect the Seafood Fishes. VGG16 performed better among the pre-trained models, with test accuracy equal to 96.54% and our CNN model SFFDNet had test accuracy equal to 98.80. Accuracy was adopted as the primary evaluation metric in this study, as the dataset was balanced across nine classes, making accuracy a reliable measure of classification effectiveness. The detailed results in terms of accuracy and loss for all models are presented in Figures 6 and 7. The accuracy results obtained from all models are presented in Table 4.

Table 4. Accuracy of all models.

Model Name	Train Accuracy	Test Accuracy
SFFDNet (proposed model)	98.71%	98.80%
VGG16	94.34%	96.54%
VGG19	93.89%	94.82%
EfficientNetB7	91.67%	89.62%
DenseNet201	86.67%	86.04%
MobileNetV2	76.12%	74.32%
Xception	71.12%	72.23%
ResNet101	64.62%	63.45%
InceptionV3	62.45%	58.39%
ResNet152	55.89%	53.70%
ResNet50	53.34%	53.34%

5.1 Efficiency Analysis

To further validate the lightweight design of SFFDNet, we conducted an efficiency comparison against

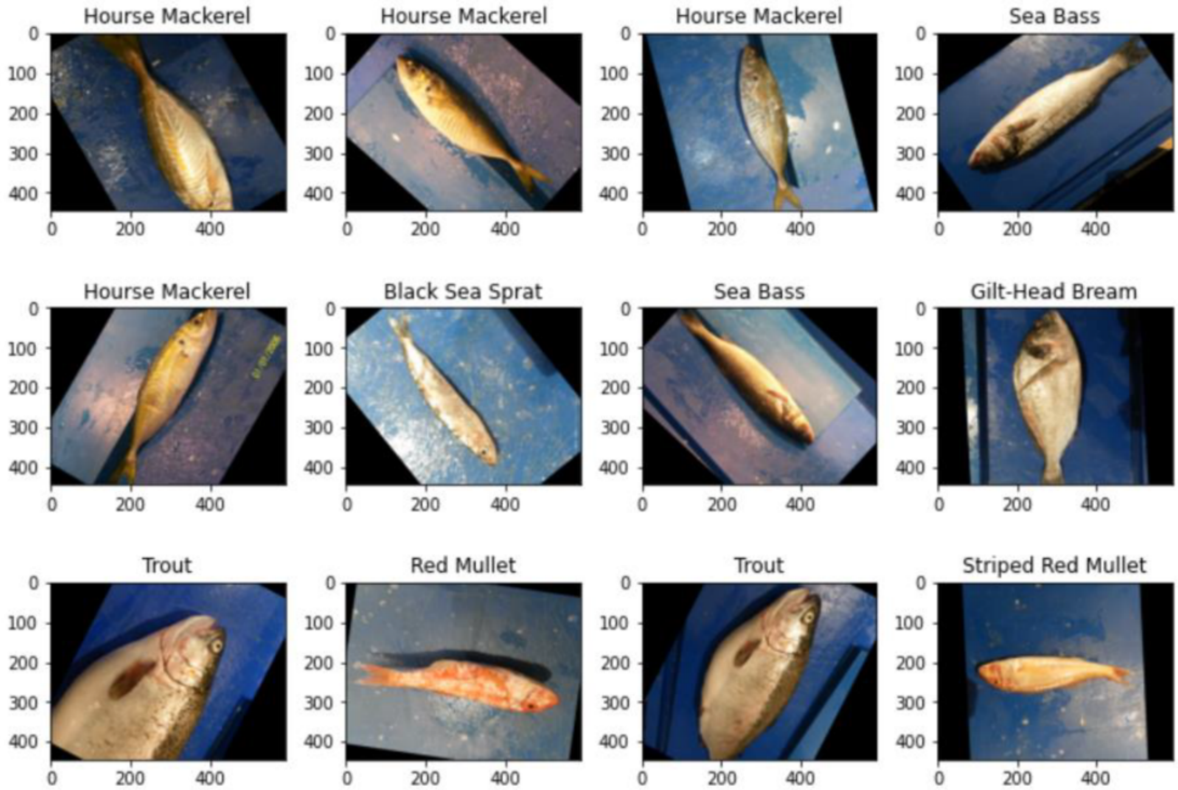


Figure 5. Representative samples from the nine fish species in the dataset (Black Sea Sprat, Gilt-Head Bream, Horse Mackerel, Red Mullet, Red Sea Bream, Sea Bass, Shrimp, Striped Red Mullet, and Trout).

representative CNN architectures. Table 5 reports approximate training time per epoch, inference latency per image, throughput, and peak GPU memory usage, measured on an NVIDIA Tesla T4 GPU with 16 GB RAM.

Table 5. Efficiency comparison of SFFDNet with baseline models.

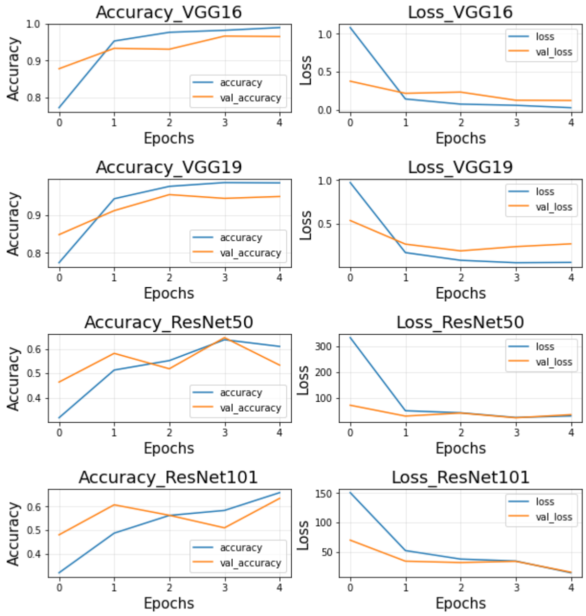
Model	Parameters (M)	Training Time/Epoch (s)	Inference Latency (ms/image)	Throughput (FPS)	Peak Memory (GB)
SFFDNet	2.49	∼45	∼3	∼330	∼0.9
VGG16	138	∼230	∼15	∼65	∼5.8
ResNet50	25.6	∼110	∼8	∼125	∼3.2
InceptionV3	23.9	∼120	∼9	∼110	∼3.5

SFFDNet consistently required significantly fewer resources than larger models. It completed training epochs approximately 5× faster than VGG16 and more than 2× faster than ResNet50. The average inference speed of SFFDNet was around 3 ms per image (330 FPS), demonstrating real-time capability. Memory consumption was also markedly lower, requiring less than 1 GB of GPU memory during inference compared to nearly 6 GB for VGG16. These findings reinforce that SFFDNet not only achieves higher accuracy but also delivers tangible computational efficiency, making it well-suited for real-time seafood inspection in constrained environments.

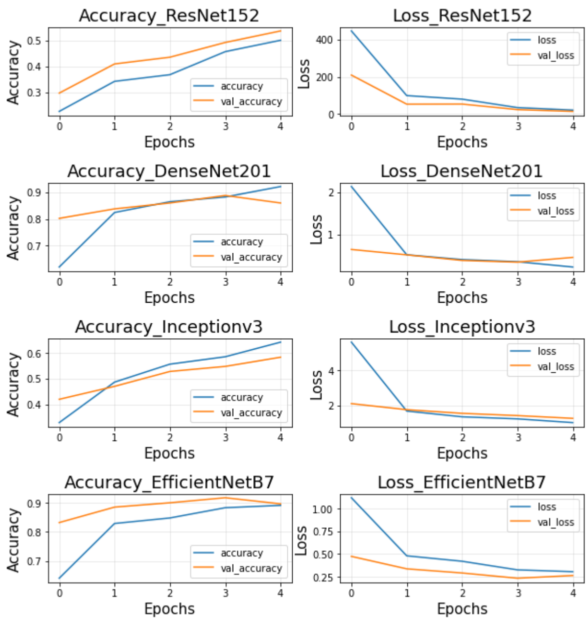
5.2 Performance Comparison

The results of the proposed method are compared with several state-of-the-art approaches, as presented in Table 6. From the comparison, it is evident that the proposed model exhibits notable performance, outperforming existing methods in terms of accuracy.

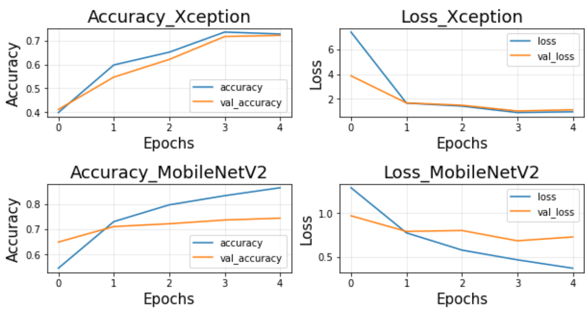
The improved performance of SFFDNet is closely tied to its architectural design. The use of hierarchical convolutional blocks with progressively increasing filters enables the network to capture both low-level textures and high-level semantic features relevant to seafood species, while keeping the overall depth manageable. The compact 19-layer structure, with only 2.49M parameters, strikes a balance between expressive power and efficiency, reducing overfitting risks associated with much deeper architectures. In addition, the integration of dropout regularization and optimized hyperparameters (batch size = 64, learning rate = 0.001, Adam optimizer) contributed to stable training and stronger generalization. By resizing inputs to 255×255 pixels, the model preserved fine-grained details, which proved critical for distinguishing visually similar species such as Red Mullet and Striped Red Mullet. Collectively, these



(a)



(b)



(c)

Figure 6. Accuracy and loss comparison of the proposed SFFDNet against ten pre-trained CNN models (VGG16, VGG19, EfficientNetB7, DenseNet201, MobileNetV2, Xception, InceptionV3, ResNet50/101/152).

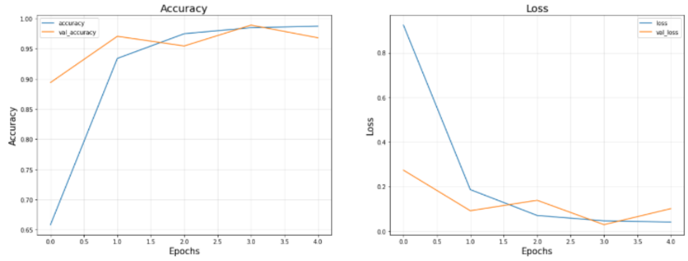


Figure 7. Training and validation accuracy and loss curves of the SFFDNet model showing smooth convergence with minimal overfitting, confirming the effectiveness of the selected hyperparameters (batch size = 32, learning rate = 0.001, Adam optimizer) and the robustness of the compact architecture.

architectural refinements explain why SFFDNet outperformed heavier CNNs like VGG16 and ResNet50, achieving both higher accuracy and lower computational cost.

These differences are meaningful for practical deployment: whereas several prior studies prioritized representational capacity at the cost of large models and heavy inference, SFFDNet balances expressivity and efficiency to deliver both improved accuracy and a much smaller computational footprint.

6 Discussion

The findings show that SFFDNet achieves both high accuracy and computational efficiency, making it a strong alternative to larger CNNs. With 98.80% accuracy and only 2.49M parameters, it outperforms deeper models like VGG16 and ResNet50 while operating at a fraction of their computational cost. This efficiency translates into faster training, lower memory use, and real-time inference speeds, which are essential in industrial seafood inspection, where throughput and reliability are critical. Importantly, SFFDNet maintained strong precision and recall across all nine classes, even when species shared similar visual features, suggesting that the architecture effectively captures subtle patterns relevant to seafood quality assessment.

Beyond accuracy, the study highlights the practical value of a domain-adapted, software-engineering-driven design. By embedding modularity, maintainability, and scalability into the architecture, SFFDNet is easier to update and integrate into existing pipelines, bridging the gap between experimental models and industrial adoption. Nonetheless, validation under more challenging real-world conditions such as imbalanced datasets,

Table 6. Comparison of proposed model with existing techniques.

Reference	Technique	Dataset	Parameters/model size	Accuracy
[20]	SegNet segmentation	A Large-Scale Dataset for Fish Segmentation and Classification	29.5M	88.69%
[21]	Hierarchical partial classifier algorithm	F4K dataset	NR	93.80%
[22]	CNN	AQUARIO28E40I dataset	NR (1–10M typical)	96%
[23]	YOLO with ResNet50	LifeCLEF 2015 from Fish4Knowledge, dataset collected by The University of Western Australia (UWA)	62M, 25.6M	91.64%, 79.80%
[24]	RF-feature point	Self-Collected	NR (tree-based)	87.30%
[25]	BP algorithm	Global Information System (GIS) on Fishes	<1M	82.25%
This study	SFFDNet (Proposed Framework)	A Large-Scale Dataset for Fish Segmentation and Classification	2.49M	98.80%

noisy images, or continuous video streams remains an important next step. Extending the framework to video analytics, multimodal data sources, and edge deployment will further strengthen its potential for large-scale, automated seafood quality monitoring.

Although SFFDNet achieved strong overall performance, the confusion matrix reveals certain limitations. A few errors were also observed for species with significant intra-class variation in pose, size, or illumination, suggesting sensitivity to visual noise and real-world imaging conditions. These cases indicate that while the model effectively captures distinctive features across most categories, additional strategies such as advanced data augmentation, multi-view inputs, or attention mechanisms may be needed to further enhance robustness for challenging species pairs. Addressing these limitations will be an important step toward ensuring reliable deployment in diverse industrial environments.

7 Limitations

In this study, accuracy was adopted as the primary evaluation metric. Given that the dataset was well balanced across nine classes, accuracy provided a reliable and interpretable measure of overall performance. This choice also enabled a consistent comparison with prior works in seafood and fish classification, where accuracy has been the dominant reporting standard. While metrics such as precision, recall, F1-score, or confusion matrices can offer additional class-level insights, our focus on accuracy ensured alignment with existing benchmarks and highlighted the superior performance of SFFDNet in both efficiency and correctness across species.

Another important consideration is the scalability of SFFDNet in industrial deployment. The current study focused on nine species, but inspection pipelines may require handling dozens of categories and even continuous video streams. The lightweight architecture of SFFDNet provides a strong foundation for such extensions, as its low computational footprint enables real-time inference. Nevertheless, scaling to larger taxonomies or video-based monitoring will demand further optimizations, including temporal modeling and incremental learning strategies. Addressing these challenges will be key to ensuring that SFFDNet can be effectively deployed in diverse seafood inspection environments at scale.

Moreover, the dataset employed in this study, comprising nine species with 1,000 images each, offered the advantage of balance, which facilitated reliable accuracy estimates. However, such balance also introduces certain constraints: it does not directly reflect the variability encountered in real-world seafood inspection, where unseen species, occlusions, or suboptimal imaging conditions (e.g., variable lighting and angles) may be common. Extending validation to more heterogeneous and industrial-scale datasets will further clarify its generalizability.

8 Conclusion

This study presented SFFDNet, a lightweight CNN framework for seafood fish detection and classification. With only 2.49M parameters, SFFDNet achieved state-of-the-art accuracy (98.80%) while requiring far fewer computational resources than models such as VGG16 and ResNet50. Beyond accuracy, the framework incorporates software engineering principles modularity, maintainability, and scalability making it practical for integration into real-time inspection pipelines and industrial environments. Despite its high performance, challenges such as noisy images, occlusions, and lighting variability remain relevant in real-world contexts. To further strengthen its practical impact, future work will focus on extending SFFDNet to video analytics for continuous monitoring, validating its robustness in operational seafood processing facilities, and exploring deployment on edge devices and embedded platforms for resource-constrained environments. In addition, establishing collaborations with industry partners will be critical to assess large-scale deployment scenarios, ensure interoperability with existing inspection systems, and accelerate the transition of SFFDNet from a research prototype to an industry-ready solution for automated seafood inspection.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Vianna, G. M., Zeller, D., & Pauly, D. (2020). Fisheries and policy implications for human nutrition. *Current Environmental Health Reports, 7*(3), 161-169. [Crossref]

[2] Kruijssen, F., Tedesco, I., Ward, A., Pincus, L., Love, D., & Thorne-Lyman, A. L. (2020). Loss and waste in fish value chains: A review of the evidence from low and middle-income countries. *Global Food Security, 26*, 100434. [Crossref]

[3] Sagar Naik, C., Mohan, C. O., Remya, S., Kishore, P., & Bindu, J. (2025). Artificial Intelligence Tools for Processing and Quality Detection of Fish and Fisheries Products. [Crossref]

[4] Kibriya, H., Rafique, R., Ahmad, W., & Adnan, S. M. (2021, January). Tomato leaf disease detection using convolution neural network. In *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)* (pp. 346-351). IEEE. [Crossref]

[5] Wang, J., Yue, H., & Zhou, Z. (2017). An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network. *Food control, 79*, 363-370. [Crossref]

[6] Tao, X., Li, H., Lu, J., & Wang, X. (2012, August). Computer vision technology for seafood quality evaluation. In *2012 International Conference on Computer Science and Service System* (pp. 1888-1891). IEEE. [Crossref]

[7] Linsen, Z., Guanglei, S., & Junbiao, D. (2006). Research on management of traceability of beef safety system through the Internet based on JSP. *Chinese Agricultural Science Bulletin, 22*(12), 13-16.

[8] Wang, Y., Zhang, H., Shi, J., Wang, S., Zhou, M., & Wang, D. (2015). Quality information traceability system based on seafood's production process. *Transactions of the Chinese Society of Agricultural Engineering, 31*(14), 264-271. [Crossref]

[9] Yan, B., Shi, P., & Huang, G. (2013). Development of traceability system of aquatic foods supply chain based on RFID and EPC internet of things. *Transactions of the Chinese Society of Agricultural Engineering, 29*(15), 172-183. [Crossref]

[10] Wang, Y., Yan, H., & Yin, Z. (2016). A quality traceability system for seafood based on bill of lots. *Revista de la Facultad de Ingeniería, 31*(4), 264-271. [Crossref]

[11] Nicolae, C. G., Moga, L. M., Bahaciu, G. V., & Marin, M. P. (2017). Traceability system structure design for fish and fish products based on supply chain actors needs. *Scientific Papers. Series D. Animal Science, 60.*

[12] Cui, S., Zhou, Y., Wang, Y., & Zhai, L. (2020). Fish detection using deep learning. *Applied Computational Intelligence and Soft Computing, 2020*(1), 3738108. [Crossref]

[13] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv:1704.04861*.

[14] Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2020). IoT, big data and artificial intelligence in agriculture and food industry. *IEEE Internet of Things Journal*, 9(9),

6305-6324. [Crossref]

[15] Rasheed, J. (2021, September). A sustainable deep learning based computationally intelligent seafood monitoring system for fish species screening. In *2021 International conference on artificial intelligence of things (ICAIoT)* (pp. 1-6). IEEE. [Crossref]

[16] Yildiz, M. B., Yasin, E. T., & Koklu, M. (2024). Fisheye freshness detection using common deep learning algorithms and machine learning methods with a developed mobile application. *European Food Research and Technology, 250*(7), 1919-1932. [Crossref]

[17] Kumaravel, B., Amutha, A. L., Mary, T. P., Agrawal, A., Singh, A., Saran, S., & Govindarajan, N. (2025). Automated seafood freshness detection and preservation analysis using machine learning and paper-based pH sensors. *Scientific Reports, 15*(1), 1-14. [Crossref]

[18] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[19] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR.

[20] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708). [Crossref]

[21] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520). [Crossref]

[22] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251-1258). [Crossref]

[23] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826). [Crossref]

[24] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778). [Crossref]

[25] Ulucan, O., Karakaya, D., & Turkan, M. (2020, October). A large-scale dataset for fish segmentation and classification. In *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-5). IEEE. [Crossref]

[26] Chuang, M. C., Hwang, J. N., & Williams, K. (2016). A feature learning and object recognition framework for underwater fish images. *IEEE Transactions on Image Processing, 25*(4), 1862-1872. [Crossref]

[27] dos Santos, A. A., & Gonçalves, W. N. (2019). Improving Pantanal fish species recognition through taxonomic ranks in convolutional neural networks. *Ecological Informatics, 53*, 100977. [Crossref]

[28] Jalal, A., Salman, A., Mian, A., Shortis, M., & Shafait, F. (2020). Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecological Informatics, 57*, 101088. [Crossref]

[29] Saitoh, T., Shibata, T., & Miyazono, T. (2016). Feature Points based Fish Image Recognition. *International Journal of Computer Information Systems & Industrial Management Applications, 8*.

[30] Alsmadi, M. K., Tayfour, M., Alkhasawneh, R. A., Badawi, U., Almarashdeh, I., & Haddad, F. (2019). Robust feature extraction methods for general fish classification. *International Journal of Electrical & Computer Engineering (2088-8708), 9*(6), 5192-5204. [Crossref]

[31] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818-2826). [Crossref]

[32] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770-778). [Crossref]

Muzzamal Hameed is currently working as an Assistant Director in the public sector, where he focuses on developing AI-driven solutions for data analysis and policy support. He holds a Master's degree in Computer Science from NUST, Islamabad, and a Bachelor's degree in Computer Science from UET Taxila. His areas of expertise include machine learning, computer vision, and deep learning. He is passionate about leveraging artificial intelligence to address real-world challenges in governance, security, and public welfare. (Email: muzzamalhameed@gmail.com)

Muhammad Haroon As a data scientist, Muhammad Haroon leverages his expertise in AI-driven solutions to drive data analysis and inform policy decisions. With a Master's degree in Computer Science, he excels in machine learning, computer vision and deep learning. Haroon is dedicated to harnessing the power of artificial intelligence to tackle pressing challenges in governance, security, and public welfare, ultimately enhancing the lives of citizens. (Email: haroon.capricorn@gmail.com)

Amna Farooq is a committed researcher specializing in artificial intelligence, healthcare analytics, and Internet of Things (IoT) environments. Her expertise spans machine and deep learning applications in medical diagnosis, speech recognition, and facial emotion analysis. She has contributed to several IoT-driven projects, including malware detection, anomaly identification, and cyberattack prediction. Amna is deeply passionate about creating AI-based solutions for early disease detection and enhancing patient outcomes. She actively engages in interdisciplinary research, working with both clinical records and image-based medical data. (Email: amnaf4916@gmail.com)

Abdul Karim Sajid Ali received his Bachelor's degree in Electronics and Communication Engineering from Osmania University, Hyderabad, and his Master's degree in Information Technology and Management from the Illinois Institute of Technology, Chicago. He is an experienced data engineer and researcher with expertise in cloud data integration, cybersecurity, and AI-driven analytics. His work spans the development of scalable ETL pipelines, secure digital identity systems, and machine learning-based threat detection. Abdul has co-authored several peer-reviewed publications in the areas of post-quantum cryptography, secure mobile communications, and AI-enhanced cybersecurity. (Email: aali62@hawk.iit.edu)