

International Journal of Communication Networks and Distributed Systems

ISSN online: 1754-3924 - ISSN print: 1754-3916

<https://www.inderscience.com/ijcnds>

The intelligent object detection framework for detecting fish from underwater images

Kalyani Peddina, Ajay Kumar Mandava

DOI: [10.1504/IJCND.S.2025.10062690](https://doi.org/10.1504/IJCND.S.2025.10062690)

Article History:

Received:	12 September 2023
Last revised:	11 January 2024
Accepted:	17 January 2024
Published online:	02 December 2024

The intelligent object detection framework for detecting fish from underwater images

Kalyani Peddina* and Ajay Kumar Mandava

Department of Electrical Electronics and Communication Engineering,
GITAM Bengaluru Campus,
NH 207, Nagadenehalli Doddaballapur, Taluk,
Bengaluru, Karnataka 561203, India
Email: 321860404003@gitam.in
Email: amandava@gitam.edu

*Corresponding author

Abstract: Marine applications heavily rely on underwater object detection, yet challenges like complex backgrounds and image quality issues impede deep learning-based detectors. Monitoring feed pellet utilisation in aquaculture is vital for efficient resource management. This study introduces a novel framework, DYNFS, merging underwater object detection and image reconstruction using YOLO-V5. Initially, we curate an underwater image dataset, refining it to remove noise, and then employ DYNFS for classification. Our approach achieves a 98.93% accuracy rate in identifying submerged feed pellets, crucial for aquaculture efficiency. However, locating pellets remains challenging due to poor image quality and small object sizes. The enhanced YOLO-V5 networks show promise in real-world aquaculture scenarios. This framework enhances underwater object detection, offering potential benefits for marine applications and aquaculture management.

Keywords: detection; image data; convolutional neural network confusion matrix; YOLO network; recurrent neural network.

Reference to this paper should be made as follows: Peddina, K. and Mandava, A.K. (2025) 'The intelligent object detection framework for detecting fish from underwater images', *Int. J. Communication Networks and Distributed Systems*, Vol. 31, No. 1, pp.63–88.

Biographical notes: Kalyani Peddina is a Research Scholar in the Department of Electrical Electronics and Communication Engineering at GITAM Bengaluru Campus. She has organised professional with 13 years of proven teaching experience, guidance and counselling skills.

Ajay Kumar Mandava received his BTech degree in Electronics and Communications Engineering from Jawaharlal Nehru Technological University, Hyderabad, India in 2006. He received the MS and PhD degree in Electrical Engineering from the University of Nevada, Las Vegas, USA in 2010 and 2013. He is currently an Associate Professor in the Department of Electrical, Electronics & Communication Engineering at the GITAM Bengaluru Campus. His research interests include image processing,

computer vision, artificial intelligence, pattern recognition and machine learning, computational intelligence, calculus of variations, partial differential equations, nondestructive testing and evaluation.

1 Introduction

The welfare of fish is now receiving more consideration in international aquaculture. By precisely measuring the quantity of fish, it is possible to ensure efficient and accurate breeding, which is essential for maintaining proper rearing circumstances and a high standard of life (Cai et al., 2020). In intertidal and subtidal marine environments, seagrasses are present and perform a variety of crucial ecological functions, including stabilising the bottom and providing foodstuff and haunt for sea invertebrate and vertebrate types (Moniruzzaman et al., 2019). To acquire underwater image data that accurately represents crabs' natural life conditions, call for a submerged camera framework with a worldwide situating framework (Cao et al., 2020). In most cases, typical data augmentation techniques work well, however, because of the unique underwater imaging environment for marine animals, they may not work as well (Huang et al., 2019). To track patterns in relative overflow, piece, size, and fish biomass in seas and new water bodies, regular monitoring of fish populations is unavoidable (Jalal et al., 2020). These applications need accurate and effectual underwater sea assessments comprising visual quality analytics (Gilbey, 2021). Underwater object identification is significant for monitoring sea life and robot navigation (Liu, 2020).

It can offer the needful information for marine robots to operate tasks like grasping. The marine exploitation as well as investigation processes increase the interest in exploring the ocean (Cong, 2019). Design coordination, physical and measurable ways of behaving, and extraction are the essential starting points for robotised fish acknowledgement. For search type counts, resident estimates, fish calculating, the research of fly-fish associations, and ecosystem monitoring, fish recognition is also crucial (Rauf et al., 2019). Improving the accuracy rate of the small targets can play a significant role in the entire recognition accuracy level (Diamant, 2019). Generally, to explore the underwater environment and underwater objects effectual devices such as remotely operated and autonomous underwater vehicles (AUVs) are utilised (Wan and Goudos, 2020). Poor light and massive noise conditions make complexity in understand the marine image or video assessments (Tanveer and Ahmad, 2022). In addition, increase this issue is the image underwater like colour deformation, scattering, and absorption make lower contrast, and reduced the colour of the image (Wang, 2019). It is increasing the complexity level in tracking, detecting, and recognising of the papers present in the submerged climate while processing the computer vision challenges (Zhou, 2022). Facilities for counting fish are crucial for assessing the biodiversity of river systems. A typical need for modern hydropower facilities is their capacity to permit unhindered passage for migrating fish species (Vasile et al., 2020).

To learn more about fish behaviour within maritime environments, fish monitoring is crucial. Researchers can learn about the condition of the marine biological systems by the

location, division, and conveyance of fish inside them (Alshdaifat et al., 2020). Several traditional models were utilised to process the underwater images yet the information on the object like locating and tracking is lost and deformed (Zhang, 2021). Underwater image enhancement benchmark (UIEB) (Li et al., 2019), effective underwater image enhancement method (Song et al., 2020), new robotic fish platform that uses reinforcement learning and a camera stabilisation system to track targets in two dimensions in real-time (Yu et al., 2020). there are some of the drawbacks like only detect one or two objects, only efficient shallow water level, high processing duration, and low accuracy in detection process etc. This makes the scope of the operation restricted, the risk rate is high and the function is complicated in manner. These reasons motivate us to develop an effectual deep learning (DL)-dependent architecture to process the underwater image and track the objects present in the underwater environments.

2 Related work

A few recent associated works are elaborated as follows:

Underwater image process techniques are the emerging method for exploring objects in the underwater environment. Due to the low optic image tool, and deformation of the image colour, the efficiency of the underwater imager detector diminishes. To solve these issues, Yeh (2021) proposed a low-weight deep underwater object identification network for studying object identification as well as colour distortion. This model effectively identifies the underwater objects compared with other models. However, there is a problem with the different sizes of the images.

Due to some drawbacks like background complexity, and image blurs the identification process of underwater objects is still a tough task. So, Sung (2020) proposed a new technique that combines real-world sonar pictures by a Generative Adversarial Network. This model can efficiently identify the underwater object by producing real visuals and also real sonar visuals with different angles and conditions of the pictures. Yet, there is no intrinsic metric evaluation.

Aquaculture in the ocean can be done more easily and with less effort thanks to computer vision technology. However, the marine fisheries' underwater picture degradation issues impede computer vision technology's ability to further understand and analyse underwater data. To overcome these issues, Wang et al. (2021) have proposed a start-to-finish coordinated double channel network model which utilises the marine snow evacuation module in light of nearby leftover learning technique to eliminate the white marine snow in the high-recurrence layer of pictures and the submerged picture upgrade module given lingering thick organisation to perform variety remedy and dehazing for the low-recurrence layer of pictures.

To learn about and comprehend the underwater world, underwater photographs are crucial. The dependability of underwater intelligent systems can be ensured by high-quality underwater photographs. Poor contrast, blurring, and low light in underwater photos have a significant impact on how underwater information is perceived and processed. Outstanding an absence of sufficient teaching records and efficient network architectures, underwater picture enhancement algorithms currently perform in an

unsatisfactory manner. To overcome these issues, Yang et al. (2020) have proposed a restrictive generative ill-disposed network where a multi-scale generator creates a clear underwater image.

Autonomous systems based on artificial intelligence networking with an active situation are required to continually learn, build, and enhance the knowledge acquired. Most artificially intelligent systems operate in an insulated knowledge paradigm and lack this capability. To overcome these issues Irfan et al. (2021) have proposed a lifelong learning model to address the complex issue of underwater image classification in the real world. As well as advancing social incorporation, dynamic citizenship, and self-improvement, it is critical for an individual's seriousness and employability.

Identification of fish species is essential for the aquaculture and fishing sectors, water body stock management, and aquatic environmental monitoring. Traditional methods for identifying fish species are expensive, inefficient, expert-based, and inappropriate for use in a great measure. To overcome these problems, Banan et al. (2020) have proposed automating the distinguishing proof of four financially huge carp species required the development and application of a real-time, non-destructive technique to attain the highest level of precision.

The narrow visual field of the camera drastically restricts the receiving of evidence for submerged optical environment perception. To overcome these issues, Qin (2020) has proposed an AUV that uses sonar and a subsurface camera to perceive the environment, as well as to process underwater images for object tracking and detection. The assimilation and scattering of underwater light severely impair underwater imaging, particularly by lowering image contrast.

Zhao et al. (2021) have introduced Compositing FishNet, a novel framework for automated fish detection and identification in underwater videos. Traditional methods struggle with poor image quality and unpredictable fish movements. Compositing FishNet utilises a composite strength and an improved path aggregation system to reduce the impact of underwater environmental factors and enhance the transmission of relevant object information. Experimental results show significant improvements, with average precision (AP) 0.5 : 0.95, AP 50, and average recall (AR) max = 10 reaching 75.2%, 92.8%, and 81.1%, respectively. The method proves effective for fish detection and Identifying entities within intricate underwater settings, such as oceans and aquaculture, poses a challenge.

Stavelin et al. (2021) explore the feasibility of employing machine learning (ML) techniques, specifically the YOLO version 3 algorithms, for fish detection in challenging conditions Marked by dim illumination, the undesired proliferation of algae, and elevated noise levels, the algorithm, which was trained using images acquired during daylight under natural lighting conditions, exhibited effective fish detection in images captured during nighttime under artificial lighting. The achieved total normal precision score was 88%. Subsequently, principle component analysis (PCA) was employed to examine the structures educated in various layers of the system. The study concludes that, for the specific application of object detection, the network can be shortened significantly as many piece detectors are found to be redundant.

Detecting underwater targets is fundamental for ensuring the autonomous operation of underwater vehicles and stands as a key technology in marine exploration.

The challenging and unique underwater environment often leads to suboptimal detection outcomes and insufficient precision. Wang et al. (2023) have developed a YOLOv5 as the overarching basic structure for the aim detection algorithm and introduced enhancements to improve detection precision in underwater conditions. The method under consideration has been subjected to testing using the dataset for underwater target detection sourced from the China Underwater Robot Professional Competition. The outcomes of the experiments indicate a detection precision (P) of 85%, a recall (R) of 82.2%, and a mean average precision (mAP) of 87.5%. The validation of the proposed approach establishes its effectiveness, showcasing superior performance in underwater target detection when contrasted with traditional models.

Al Muksit et al. (2022) presented a YOLO-Fish, a DL model for fish detection, comprising two models, YOLO-Fish-1 and YOLO-Fish-2. YOLO-Fish-1 improves YOLOv3 by addressing up-sampling step size issues, minimising the misdetection of small fish. YOLO-Fish-2 enhances the typical by incorporating Spatial Pyramid Pooling for dynamic environment adaptability. The evaluation involves two datasets: DeepFish (15k bounding box observations across 4505 images) and OzFish (43k bounding box observations across 1800 images). YOLO-Fish-1 and YOLO-Fish-2 achieve average precision of 76.56% and 75.70%, surpassing YOLOv3 in unconstrained marine environments. Despite being lightweight compared to YOLOv4, both models exhibit comparable performance.

Utilising computer vision technology in aquaculture holds the potential to enhance fish detection efficiency, monitor health, and optimise overall aquaculture management for increased profit. Traditional keypoint detection methods have demonstrated drawbacks, including in underwater environments, there is a notable lack of efficiency, subpar accuracy, and limited robustness. In response to these challenges, Dong et al. (2023) have introduced a novel approach Built upon object detection and point regression models, the method focuses on the precise localisation of fish key points. The network neck has been enhanced to elevate the accuracy of fish detection. Empirical findings affirm the efficacy of the proposed approach in successfully identifying individual underwater fish and precisely estimating key points. Additionally, the source code and labelled datasets for fish detection and keypoint location are accessible for further use.

The challenging underwater environment often leads to subpar detection results and low precision. Wang et al. (2023) developed the YOLOv5 framework serves as the foundational structure for the target detection algorithm. Improvements are implemented to enhance finding correctness in underwater settings. A novel devotion mechanism, channel and spatial fusion attention (CSFA), is introduced and integrated into the YOLOv5 network framework. This addition enables the network to concentrate on both the prominent features and spatial information of the detected object. The overall recall (R) achieved 82.2%, and the mean average precision (mAP) reached 87.5%. The planned method's efficacy was validated, demonstrating superior underwater target detection concert compared to conventional models.

Fish play a crucial role in both marine and freshwater environments, and understanding their morphological and computational aspects, encompassing dimensions, form, and placement, is essential for fisheries and fish observation. Conventional methods for acquiring fish details are often manual and inefficient. These approaches are

time-consuming, less precise, and demand significant resources. To address these challenges, two innovative networks, the parallel feature Haider et al. (2023) developed a fusion-based segmentation network, and the progressive information fusion-based segmentation network (PIFS-Net), were established for pixel-wise fish segmentation. Acting as the foundational model, PFFS-Net utilises parallel feature fusion to enhance segmentation performance. The ultimate PIFS-Net model incorporates a progressive spatial feature fusion (SFF) mechanism aimed at enhancing segmentation accuracy. Automatic and accurate segmentation of fish represents a significant advancement toward building intelligent aquatic ecosystems.

Enhancing the ability to discern diverse fish species holds significant importance in the field of aquaculture. This capability is crucial for safeguarding fish populations, monitoring their health, and understanding their nutritional systems. However, traditional ML methods struggle to identify objects in pictures featuring intricate backgrounds, particularly in low-light conditions. Hamzaoui et al. (2023) have developed to enhance the FishDETECT model, an improved version utilising transfer learning with the pre-trained FishMask model for fish recognition and classification, which exhibits superior performance compared to the conventional YOLO v5 model. Through rigorous testing in diverse and intricate scenes, our experimental results reveal that FishDETECT achieves higher accuracy rates. Precision, Recall, and mAP50 evaluation metrics showcase impressive values of 0.962, 0.978, and 0.995, respectively, highlighting the enhanced capabilities of our novel model.

The challenges of the existing mechanisms are mentioned in the Table 1.

Table 1 Challenges of the current works

<i>S. No.</i>	<i>Authors name</i>	<i>Method</i>	<i>Advantage</i>	<i>Disadvantage</i>
1	Yeh (2021)	Object identification network	It is simple and quick to save and retrieve large, complex datasets	When viewed from several angles, an object can appear radically different
2	Sung (2020)	Generative adversarial network	GANs produce data that resembles real data is one of their best features	It could lead to erratic and sluggish training
3	Wang et al. (2021)	Start to finish incorporated double channel network model	Data transmission rates through integrated services digital network (ISDN) are substantially faster	High cost
4	Yang et al. (2020)	Conditional generative adversarial network (CGAN)	Efficient network architectures	Poor contrast, blurring, and low light in underwater photos
5	Irfan et al. (2021)	Lifelong learning model	Competitiveness and employability	Complex issue of underwater image classification in the real world

Table 1 Challenges of the current works (continued)

<i>S. No.</i>	<i>Authors name</i>	<i>Method</i>	<i>Advantage</i>	<i>Disadvantage</i>
6	Banan et al. (2020)	Automating the identification of four economically significant carp species	Non-destructive technique attained the highest level of precision	Inappropriate for use on a large scale
7	Qin (2020)	Autonomous underwater vehicle (AUV)	Mapping and tracking of chemosynthetic ecosystems and low-temperature fluid escape characteristics	Due to auditory interference, they are also less appropriate for densely populated locations
8	Zhao et al. (2021)	FishNet	Effective for fish detection and identification	Poor image quality and unpredictable fish movements
9	Stavelin et al. (2021)	You Only Look Once version 3	The overall average precision score was 88%	Dim lighting, unwelcome proliferation of algae, and elevated levels of noise
10	Wang et al. (2023)	YOLOv5	Improve detection precision in underwater conditions	Insufficient precision
11	Al Muksit et al. (2022)	YOLO-Fish, a deep-learning model	Achieve average precision of 76.56% and 75.70%	Minimising misdetection
12	Dong et al. (2023)	Regression models	Enhance fish detection efficiency	Detection accuracy is less
13	Wang et al. (2023)	YOLOv5	Overall recall $\text{\textcircled{R}}$ achieved 82.2%	Low precision
14	Haider et al. (2023)	Fusion-based segmentation network (PFFS-Net) and the progressive information fusion-based segmentation network (PIFS-Net)	Improve segmentation accuracy	Computational cost is high
15	Hamzaoui et al. (2023)	YOLO v5	Improved readability	Redundancy

The key role of this current investigation work is described as trails

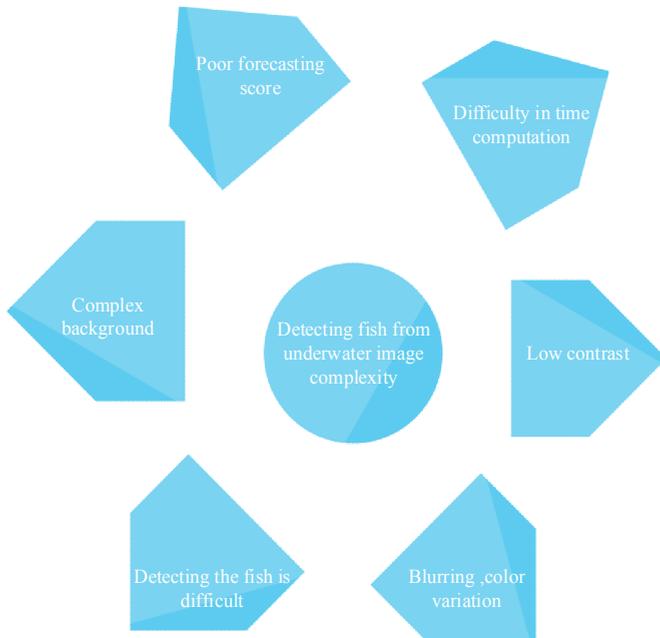
- Firstly, the underwater image dataset has been considered it has included fish and different objects.
- Then a novel DYNFS was planned with essential cleaning and the detection limitation.
- Primarily, the noise was filtered and the refined records were imported to the organisation phase of the DYNFS system.
- Here, the incorporation of dove fitness in the Yolo network has afforded the finest fish detection outcome.

- Subsequently, the key detection parameters were validated and compared with other conventional models like Inception, Inception ResNet V2, CNN- SENet, You Only Look Once (Yolo)V5, Yolo V5-KF, Yolo V5-PA, KAYOLO, CAXCF-LL, DenseNet, Residual Neural Network (ResNet), MobileNet, Xception in terms of exactness, recall, F-score, accuracy, and blunder rate.

3 System typical and problem statement

The detection application was most attracted by the use of neural strategies. Figure 1 depicts the system model for the challenges in fish detection from underwater image data.

Figure 1 System model of detecting fish from underwater image data (see online version for colours)



However, if the images are too complex or in different featured pixels then it has recorded a poor forecasting score. Low contrast, blurring, colour variation, and obscuration are common issues with underwater photos, which result in poor quality and make it challenging to find and identify fish in large schools (Cotter and Staines, 2023). Here, the underwater object data in the image format was considered as the dataset for testing the proposed framework. Hence, the underwater image data contains several objects from which the fish object was taken into consideration in this present research work. So, from the several objects in the image data, detecting the fish is difficult (Fu et al., 2023). Considering these demerits, the present research has planned to implement a novel optimised deep network system for the detection objective. Also, the application that has been considered to check the detection performance is underwater image data.

4 Proposed methodology

A novel Dove-based Yolo Neural Forecasting System (DYNFS) was introduced in this present research work for detecting fish underwater. Here, the underwater object detection database was considered then a novel DYNFS was executed for predicting the present object from the image underwater data.

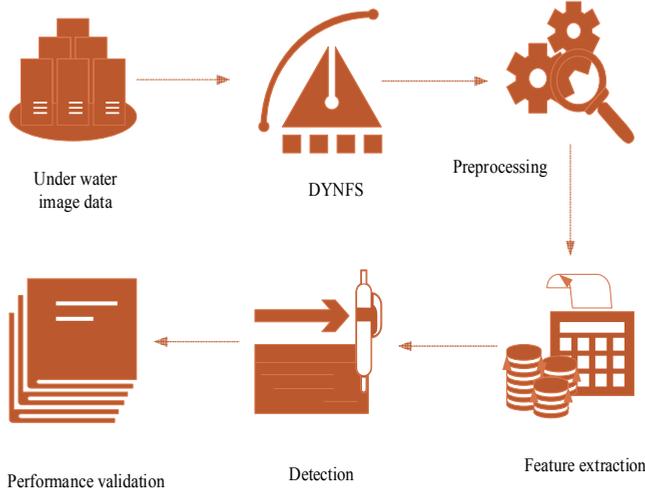
Initially, the trained images were preprocessed and imported to the classification frame. In that feature extraction, the detection process has been performed with the help of the dove fitness function. Here, the classification layer of the Yolo is tuned by the dove fitness that has afforded the finest outcome. The fitness function is critical and depends on the specific goals and characteristics of the problem being addressed. The fitness function is used for various metrics, such as precision, recall, F1 score, or other domain-specific criteria.

In the context of object detection, the dove fitness function can impact the performance of the system in the training process. Dove fitness function helps guide the model towards learning relevant features and making accurate predictions. The impact on performance varies depending on the characteristics of the data, the complexity of the problem, and the suitability of the fitness function for the specific task at hand. The proposed procedure is diagrammatically represented in Figure 2. Initially, the underwater image dataset has been considered to include fish and different objects. Then a novel DYNFS was considered with mandatory filtering and the detection parameter. Primarily, the noise was filtered and the refined data was introduced to the classification phase of the DYNFS system. The impact of these filtering techniques on detection accuracy depends on the definite features of the images and the nature of the noise. While, the filtering method enhances image quality and makes it easier for object detection algorithms to identify fish, to evaluate the chosen techniques to ensure that they do not inadvertently remove relevant features or introduce artefacts. Ultimately, the effectiveness of noise reduction techniques in improving fish detection accuracy should be validated through thorough experimentation and evaluation using appropriate performance metrics. Here, the incorporation of dove fitness in the Yolo network has afforded the finest fish detection outcome. Subsequently, the key detection parameters were validated and compared with other conventional models like Inception, Inception ResNet V2, CNN-SENet, You Only Look Once (Yolo)V5, Yolo V5-KF, Yolo V5-PA, KAYOLO, CAXCF-LL, DenseNet, Residual Neural Network (ResNet), MobileNet, Xception in terms of, accuracy, recall, F-score, precision, and error rate.

YOLO is designed for speed and efficiency, making it suitable for deployment on resource-constrained devices or in scenarios where processing time is critical. This efficiency is achieved through the use of a sole neural network for the direct prediction of bounding boxes and class probabilities, as opposed to region-based approaches that involve multiple stages. The choice of a neural network framework should also consider the specific requirements of the application, available computational resources, and the characteristics of the dataset being used for training. Additionally, newer versions of YOLO addressed some limitations of earlier versions. YOLO has shown robust performance even in the presence of class imbalance, which is common in many real-world datasets, including those related to underwater environments. This robustness can be critical when dealing with diverse marine life populations. YOLO tends to provide accurate object localisation, meaning it can precisely determine the location of objects

within an image. This is crucial for tasks like fish detection in underwater images, where precise localisation is necessary for monitoring and research purposes.

Figure 2 Proposed architecture (see online version for colours)



Underwater: The initial step involves acquiring an underwater image dataset containing various objects, including fish. This dataset serves as the foundation for training and testing the intelligent object detection framework. Water absorbs and scatters light, affecting colour and contrast in images. The farther the light travels through water, the more it is absorbed, leading to a loss of colours and reduced visibility. Underwater visibility can vary greatly depending on factors like water quality, presence of particles, and depth. Murky water can reduce visibility, making it challenging to capture clear images.

Preprocessing: The preprocessing phase of an intelligent object detection framework for detecting fish in underwater images involves several key steps to enhance the quality and relevance of the input data. These steps aim to optimise the images before feeding them into the object detection model. Resize the images to a consistent resolution. This ensures uniformity in input dimensions and helps in efficient processing. Normalise the pixel values of the images to a common scale (e.g., between 0 and 1). This step aids in convergence during training and ensures that the model is not sensitive to variations in pixel intensity. Remove any irrelevant or distracting elements from the images, such as artefacts or debris that could potentially interfere with the object detection process.

Feature extraction: The feature extraction phase involves capturing relevant information from the preprocessed images to represent them in a form that facilitates the detection of fish. Commonly, this is achieved through the use of convolutional neural networks (CNNs) or other deep learning architectures designed for image recognition (Dawson et al., 2023). The feature extraction phase transforms the preprocessed images into a format that allows the intelligent object-finding framework to identify and localise fish accurately in underwater environments.

Detection: The detection phase of an intelligent object detection framework for identifying fish from underwater images involves applying the trained model to new, unseen data to locate and classify fish instances within the images. To develop an underwater fish detection system, employ the TensorFlow Object Detection Application Programming Interface (API) with a pre-trained model such as Faster R-CNN or Single Shot Multibox Detector (SSD), fine-tuning it on a labelled dataset of underwater images (Tasyurek and Arslan, 2023). Tackle challenges unique to underwater environments through data pre-processing, addressing issues like lighting variations and water distortions. Set up the API, configure the model for training, and conduct the training process while adjusting parameters as needed. Evaluate the model, export it for inference, and seamlessly integrate it into applications using TensorFlow Lite or TensorFlow Java API, considering resource constraints. This comprehensive approach ensures the creation of a robust and efficient underwater fish detection framework, ready for real-world deployment.

4.1 Design of DYNFS

The recommended approach makes use of a neural network with dove swarm optimisation. The proposed DYNFS 5 layers: An input layer, a Hidden layer, a Feature extraction layer, an optimal layer, Output layer. A brain organisation, a sort of computerised reasoning methodology, guides PCs to break down information in a way demonstrated by the human mind. Consequences be damned is a strategy that gives constant item identification utilising brain organisations. The prominence of this calculation is because of its exactness and snappiness. It has been applied in different ways to distinguish creatures, people, stopping meters, and traffic signals. A computational strategy tries to upgrade a competitor arrangement iteratively for a predefined quality measurement to enhance an issue. The image data is imported into the input layer. The hidden layer follows, which disregards the incorrect value. The desired data is extracted during feature extraction. Dove swarm optimisation's fitness function is inserted in the optimal layer and provides the outcomes from the output layer. Layers of DYNFS are shown in Figure 3.

Feature representation: Deep neural networks excel at automatically learning hierarchical representations of features from raw data. In the context of underwater images, which often contain complex patterns and textures, deep networks can effectively capture intricate details and patterns relevant to fish detection. This ability to automatically learn hierarchical features is crucial for handling the diversity of underwater scenes (Yuan et al., 2023).

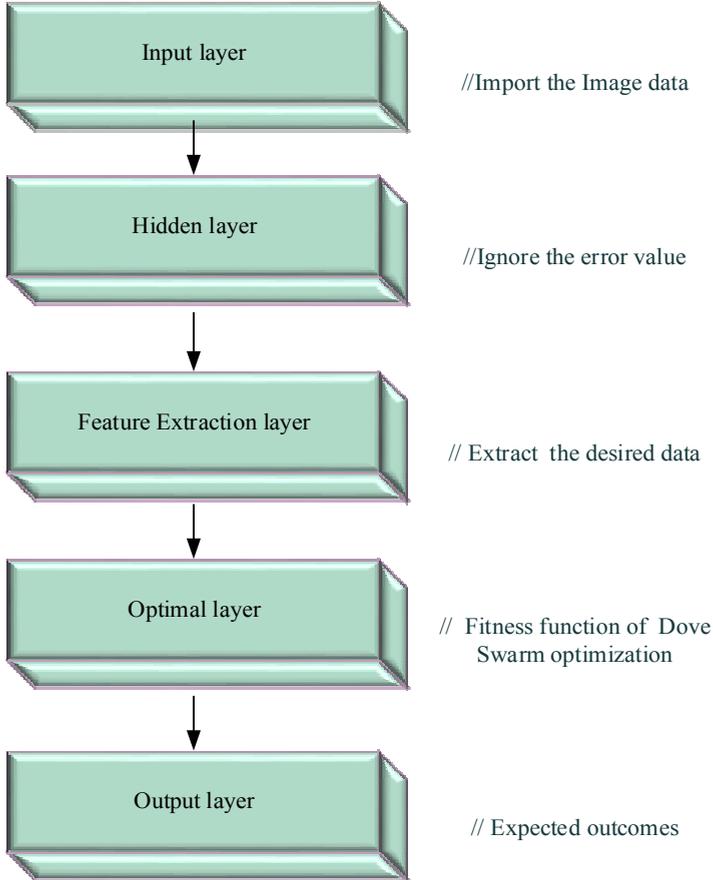
Non-linearity and complexity: Underwater environments exhibit a high degree of variability, including changes in lighting conditions, water clarity, and the presence of vegetation or other marine life. Deep networks are capable of modelling complex, non-linear relationships within the data, enabling them to adapt to the diverse and dynamic nature of underwater imagery.

End-to-end learning: Deep networks enable end-to-end learning, allowing the model to directly map raw input data to achieve fish detection in the desired output without relying on manual feature engineering. This is advantageous in scenarios where the underlying

patterns and features that define fish in underwater images might be challenging for human experts to articulate explicitly (Paraschiv et al., 2022).

Transfer learning: Pre-trained deep neural network architectures, such as those trained on large-scale image datasets like ImageNet, can be fine-tuned for specific tasks with relatively limited labelled data. Transfer learning leverages the knowledge acquired during pre-training, facilitating faster and more effective training on the target task of fish detection in underwater images (Heidenreich et al., 2023).

Figure 3 Layers of DYNFS (see online version for colours)



4.2 Data initialisation

Initialisation is the process of giving a data object or variable a starting value. The data initialising equation (1) is given by,

$$H(F_g) = \varphi(f_1, f_2, f_3, \dots, f_n) \quad (1)$$

where, F_g is the dataset, φ is the data function, f_n is the data element.

4.3 Hidden layer

An activation function is employed to generate output from a set of weighted inputs. It also neglects the error values and detects the images of the proposed DYNFS model. The hidden layer of DYNFS is a key benefit because it ignores undesired data. The equation (2) is given by,

$$H(F_g) = [\sigma(m) - \mu(m)] \quad (2)$$

where F_g the dataset is, $\mu(m)$ is the error value.

Activation functions are typically associated with the hidden layers of neural networks, and they introduce non-linearity to the model, enabling it to learn complex patterns. In the context of YOLO or DYNFS, the activation function used in the hidden layers would likely play a crucial role in allowing the neural network to capture intricate relationships within the data. Here, the rectified linear unit (ReLU) and Softmax activation function are utilised in neural networks. Softmax is frequently employed in the output layer when dealing with multi-class classification problems. This function converts the raw output scores of the model into probability distributions across various classes.

4.4 Feature extraction

The most common way of transforming crude information into mathematical highlights that might be handled while keeping the data in the first informational index is alluded to as component extraction. The undesirable 741 characteristics are removed by equation (3) is given as.

$$H(F_{g+1}) = F_g - D_g \quad (3)$$

where F_g is the wanted data, D_g is the unwanted data.

4.5 Classification

The category with the highest output value is chosen by the classification network. Classification neural networks become quite powerful when used in a hybrid system with different kinds of predictive neural networks. Our suggested model DYNFS, is presented and detects the fish in the classification layer.

$$H(CL) = \begin{cases} \text{If } A_d = 1 & \text{Fish detect} \\ \text{If } A_d < 1 & \text{Fish is not detected} \end{cases} \quad (4)$$

where H is the classification function, CL is the classification, A_d detect the fish.

```

Algorithm 1: Proposed DYNFS Method

Start
    {
        {
            Initialize the data
            Int  $H(F_g) = \varphi(f_1, f_2, f_3, \dots, f_n)$ 
        }
    }

Pre-processing
    {
        int  $H, \mu, \sigma$ 
        // Initializing the preprocessing variables
         $H(F_g) = [\sigma(m) - \mu(m)]$ 
        // Neglecting the unwanted variables
    }

Feature extraction
    {
        int  $H(F_g)$ 
        // initialize the feature variables
         $H(F_{g+1}) = F_g - D_g$ 
        // wanted features are extracted
    }

Optimal layer
    {
        {
            If  $A_d = 1$ 
            //if the support value is 1, then the fish is Detected
        }
        Else
        }
    }

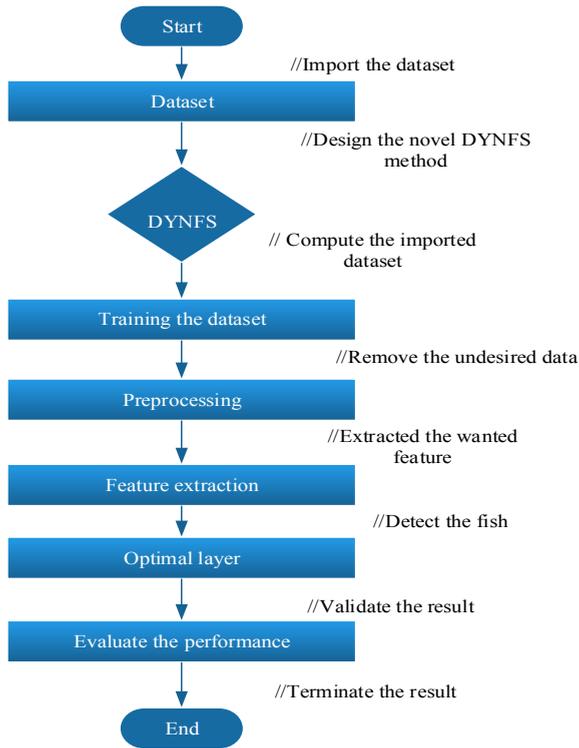
}
End

```

The workflow of the proposed DYNFS is illustrated in Figure 4. The algorithm of the presented model is represented in pseudocode format in the algorithm.

5 Result and discussions

The suggested method is also detectable and recreates the original datasets. These strategies enable the dataset's data to be detected. Here, the underwater object detection database was taken into account to forecast the current object from the underwater image data. In the classification layer, the Dove swarm algorithm's fitness is updated.

Figure 4 Workflow of DYNFS proposed (see online version for colours)

5.1 Experimental system

The proposed model is executed by using a Python, 8 GB RAM, Windows 10 computer, and Intel I5 processor.

5.2 Dataset description

The collection includes seven types of aquatic species, each with its own set of box locations. The seven species are ‘fish’, ‘jellyfish’, ‘penguin’, ‘puffin’, ‘shark’, ‘starfish’, ‘stingray’. The training, validation, and test sets of the dataset have previously been separated. It has 638 photos with annotations in the YOLO v5 PyTorch format for the creatures. Automatic pixel data rotation (with EXIF-orientation stripping) Scale down to 1024×1024 (Fit within). This dataset comprises 638 images collected by Roboflow from two aquariums located in US: The Henry Doorly Zoo in Omaha on October 16, 2020, and the National Aquarium in Baltimore on November 14, 2020. The Roboflow team, with some assistance from SageMaker Ground Truth, meticulously labelled the images and organised them into 7 classes for object detection. The dataset is available in various popular formats, facilitating the training of ML models. Potential applications of this dataset include coral reef conservation, environmental health monitoring, swimmer safety, pet analytics, automated feeding, and more. To facilitate ML experiments, the dataset undergoes a split operation, resulting in training and testing subsets. The split ratio is set at 70% for training and 30% for testing, aligning with common practices in

ML experiments. The resulting subsets are denoted as ‘train’ and ‘train’ for training data, while ‘test’ and ‘test’ represent the testing data. Ensure the dataset includes a wide variety of fish species, poses, lighting conditions, and underwater environments. Adjust learning rates, batch sizes, and other hyperparameters during training to achieve optimal convergence. Apply data augmentation techniques to increase the variability of the training dataset and enhance model generalisation.

5.3 Performance metrics

This segment examines the exhibition of the proposed strategy and the experimental setup. Several metrics, including Accuracy, Recall, Precision, F-score, and Error rate are used to assess the efficiency of the framework for detecting fish from underwater image data.

5.3.1 Accuracy

It is the level of exactness with which the framework model works. By and large, it is the extent of accurately anticipated recognition of all perceptions. The precision is determined by utilising condition (5),

$$Accuracy(\%) = \frac{H_S + H_N}{H_S + H_N + B_S + B_N} \quad (5)$$

In this case, H_S and H_N display the false positive and false negative values, whereas B_S and B_N show the true positive and true negative values, respectively.

5.3.2 Recall

The review is determined as the extent of valid – positive figures to add up to genuine up-sides and bogus negatives. It indicates the fraction of estimates that were successfully identified as cancer using equation (6),

$$Recall(\%) = \frac{H_S}{H_S + B_N} \quad (6)$$

Here, H_S is the false positive rate, B_N true negative rate.

5.3.3 Precision

Precision is evaluated as the number of correct positive estimates alienated by the overall positive estimates. The fraction of the exact diagnosis of the disease in the exaggerated region is determined using equation (7),

$$Precision(\%) = \frac{H_S}{H_S + B_S} \quad (7)$$

In this case, H_S is the false positive rate, B_S or true positive rate.

5.3.4 F-score

The f-measure is typically used to calculate the system's accuracy based on the input data. The f-score is employed to assess the binary system classifications. Precision and recall metrics are used to evaluate the F-measure. Equation (8) is a representation of the f-measure equation.

$$Fscore(\%) = 2 \times \left(\frac{M_r \times N_c}{N_c + M_r} \right) \quad (8)$$

where M_r is the precision, and N_c is the recall.

5.3.5 Error rate

The expression 'mistake rate' alludes to an estimation of the extent of a model's made as for the genuine expectation blunder. Error rate = 1 - accuracy, the equation is given by,

$$Errorrate = 1 - \left(\frac{H_s + H_N}{H_s + H_N + B_s + B_N} \right) \quad (9)$$

5.4 Confusion matrix

The normal results for a characterisation task are summed up in a disarray network. Each class portion of right and inaccurate expectations is added up to separate utilising count values. The disarray network's key can be viewed here. A disarray grid is a table where the genuine qualities address how well a grouping model performed utilising a bunch of test information. The quantity of species in each genuine class is addressed by the lattice's lines, while the quantity of species in each anticipated class is addressed by its sections. The true positive (TP), true negative (TN), false positive (FP), and False negative (FN) instances of the partaking classifications of classes are additionally referenced in the disarray framework and are used to work out significant attributes like accuracy and recall. Diagonal elements in the confusion matrix of our planned model are illustrated in Figure 5.

In the disarray grid, corner-to-corner parts address certified up-sides (the model predicts the positive class precisely), though any remaining things aside from askew components address misleading up-sides (the model erroneously predicts the positive class). For every one of the 7 characterisations that were analysed in the disarray network, the power of the blue tone becomes as the quantity of tests increments, and a scale on the chart's right side shows the number of tests for each class.

5.5 Experimental results

This section outlines the investigational outcomes of approaches for detecting fish from underwater image data. Figure 6 depicts the output for the proposed DYNFS. Figure 6(a) shows the input images, whereas the output image is exposed in Figure 6(b). Figure 7 displays the proposed method of train, testing, and validation.

- *Training phase:* In this stage, the framework utilises a labelled dataset, such as the one described earlier, which includes images from aquariums with corresponding class labels. The model is trained on the ‘Xtrain’ dataset, where ‘Xtrain’ denotes the input images, and ‘Ytrain’ represents the corresponding labels. The intelligent object detection algorithm learns to recognise patterns and features indicative of fish within the underwater images.
- *Testing phase:* The trained model is then evaluated using the ‘Ytest’ dataset, where ‘Xtest’ comprises new and unseen images not used during the training phase. The goal is to assess the model’s generalisation and its ability to accurately detect fish in diverse underwater conditions. Performance metrics such as recall, F1 score, and precision may be employed to quantitatively evaluate the model’s effectiveness.
- *Validation phase:* The validation phase involves a separate set of data, ‘validate’ and ‘validate’, which is distinct from both the training and testing datasets. This subset serves as an additional measure to ensure the model’s robustness and generalisability. Adjustments to the model architecture or hyperparameters can be made based on the validation results to enhance performance.

Figure 5 Proposed model confusion matrix (see online version for colours)

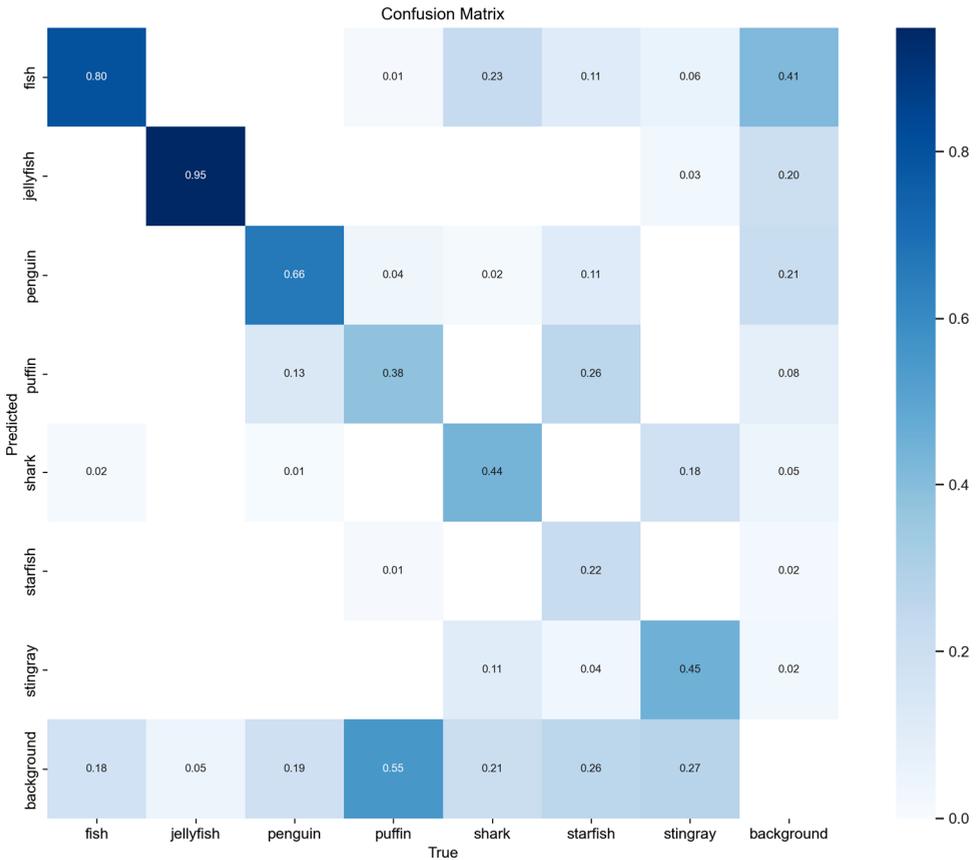


Figure 7 The proposed method of DYNFS in terms of train, testing, and validation (see online version for colours)

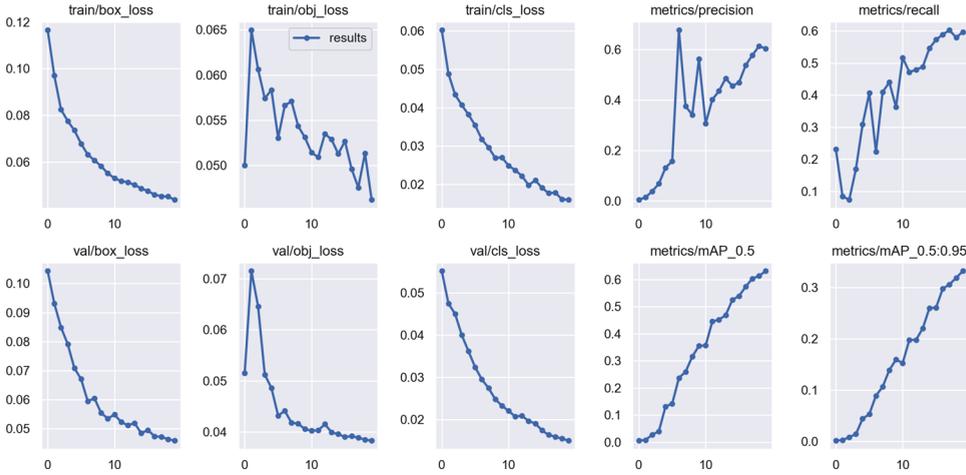
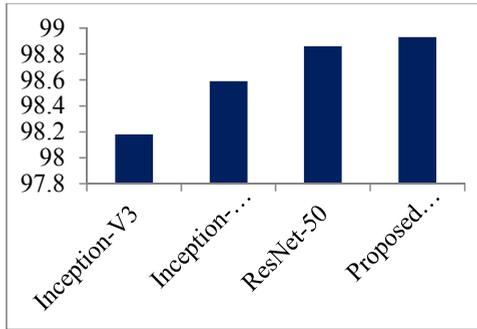


Figure 8 Comparison of the proposed DYNFS in terms of accuracy (see online version for colours)

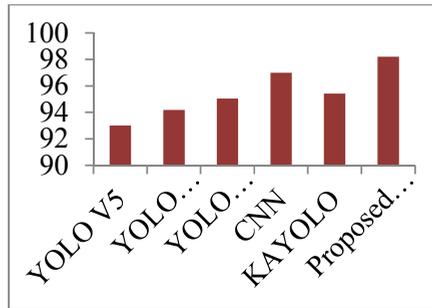


When compared to other methods like Inception-V3, and CNN-SENet the high accuracy level of ResNet-50 is 98.86%. The accuracy level of Inception-ResNet V2 is 98.59% it has a high value compared to all other existing techniques except ResNet-50. CNN-SENet has an accuracy level of 99.27%. Compared to all other approaches currently in use, it is of little utility. The proposed model has a high accuracy level; the value is 98.93%. Therefore the proposed DYNFS performs more accurately than the other current approaches as a result.

5.6.2 Comparison of the proposed DYNFS in relation of precision

Figure 9 compares the suggested DYNFS and existing approaches in terms of precision. The proposed model’s precision value for the DYNFS is 98.2%. When related to other current approaches the precision value of the proposed DYNFS is higher. In the existing method of YoLo V5, the precision value is 93.01% while comparing the existing methods of YoLo V5-KF, YoLoV5-PA, KAYOLO, and CNN the precision value is low. The YoLo V5-KF precision value is 94.17% it has a low precision value compared to CNN, YoLoV5-PA, and KAYOLO existing techniques.

Figure 9 Comparison of the planned DYNFS in terms of precision (see online version for colours)

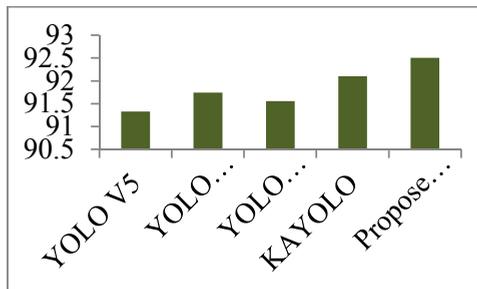


The precision rate for the YoLoV5-PA technique currently in use is 95.05%. The precision value is low when compared to the CNN and KAYOLO approaches. The precision value of KAYOLO is 95.43% it has a low value when compared to other existing techniques of CNN. The precision level of CNN is 97% it's high compared to other techniques like YoLo V5, YoLo V5-KF, YoLoV5-PA, KAYOLO, and CNN. Therefore our proposed method DYNFS has a high precision value compared to other techniques.

5.6.3 Comparison of the proposed DYNFS in standings of recall

Figure 10 compares the proposed DYNFS with current techniques in terms of recall. The proposed model's recall value for the DYNFS is 92.5%. When associated with other existing systems the recall value of the proposed DYNFS is higher. The existing method of YoLo V5 has a low recall value is 91.33% when related to all other existing systems of YoLo V5-KF, YoLo V5-PA, and KAYOLO. The YoLo V5-KF recall value is 91.74% it has a high recall value compared to the existing methods of YoLo V5, and YoLo V5-PA. The recall rate for the YoLo V5-PA technique currently in use is 91.56%. The recall value is low when compared to other available approaches YoLo V5, YoLo V5-KF. The Recall value of KAYOLO is 92.1% it has a high value when compared to other existing techniques of YoLo V5-KF, YoLo V5-PA, and YoLo V5. The Recall level of DYNFS is 92.5% its high value compared to all other techniques Therefore our proposed method DYNFS has a high recall value compared to other techniques.

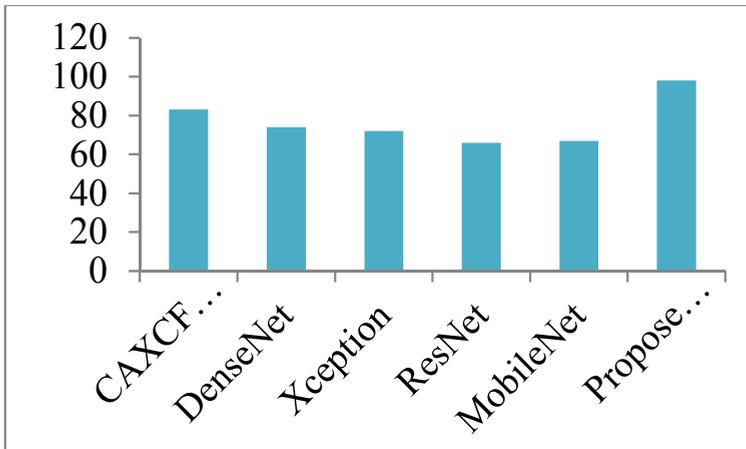
Figure 10 Evaluation of the proposed DYNFS in terms of recall (see online version for colours)



5.6.4 Comparison of the proposed DYNFS in terms of F1 score

Figure 11 displays the examination of the proposed DYNFS F1 score. The suggested DYNFS F1 score rating is 98%. While comparing the suggested DYNFS to other existing methods, the F1 score level is greater. The F1 score rate for the CAXCF-LL method currently in use is 83%. The F1 score level is high when compared to other existing methods like DenseNet, Xception, ResNet, and MobileNet. DenseNet F1 score level is 74%, which is a high value when compared to other methods like Xception, ResNet, and MobileNet. When compared to other methods like ResNet, and MobileNet the high F1 score level of Xception is 72%. The score level of ResNet is 66% it has a low value compared to all other existing techniques like CAXCF-LL, DenseNet, and Xception. ResNet has an F1 score level of 66% it has a low value when compared to all other approaches like DenseNet, Xception, and CAXCF-LL. The projected model has a high F1 score level; the value is 98%. As a result, the suggested DYNFS executes better than the other present techniques in terms of F1 score.

Figure 11 Association of the proposed DYNFS in terms of F1 score (see online version for colours)



6 Conclusion

This paper offers better consequences are damned V5 organisation to recognise uneaten feed pellets in submerged photographs, conquering the hardships of bad quality submerged pictures and tiny focuses in feed pellet distinguishing proof. The component maps are changed using thick associations and a de-overt repetitiveness process, which enhances the You Only Look Once version 5 (YOLO-V5) network. Additionally, the enhanced network decreased the error rate by about 0.0107. The high accuracy of the proposed model is 98.93% and also increases the recall, precision, and F1 scores the

increase values are 92.5%, 98.2%, and 98. This shows that utilising a thick association and better component map extraction network is successful for expanding discovery exactness while utilising the de-overt repetitiveness strategy is compelling for speeding up. As a result, the suggested method can be employed in a real aquaculture scenario to detect submerged feed pellets. Additionally, the first stage in scientific feeding is just the detection of uneaten feed pellets. Future research can focus on collecting and annotating more underwater images from different regions, depths, and environmental conditions. This expanded dataset would encompass a broader range of fish species, including rare and less-documented ones, enabling the model to generalise better and improve overall performance.

References

- Al Muksit, A., Hasan, F., Emon, M.F.H.B., Haque, M.R., Anwary, A.R. and Shatabda, S. (2022) 'YOLO-fish: a robust fish detection model to detect fish in a realistic underwater environment', *Ecological Informatics*, Vol. 72, pp.101847.
- Alshdaifat, N.F.F., Talib, A.Z. and Osman, M.A. (2020) 'Improved deep learning framework for fish segmentation in underwater videos', *Ecological Informatics*, Vol. 59, pp.101–121.
- Banan, A., Nasiri, A. and Taheri-Garavand, A. (2020) 'Deep learning-based appearance features extraction for automated carp species identification', *Aquacultural Engineering*, Vol. 89, pp.102–053.
- Cai, K., Miao, X., Wang, W., Pang, H., Liu, Y. and Song, J. (2020) 'A modified YOLOv3 model for fish detection based on mobileNetv1 as the backbone', *Aquacultural Engineering*, Vol. 91, pp.102–117.
- Cao, S., Zhao, D., Liu, X. and Sun, Y. (2020) 'Real-time robust detector for underwater live crabs based on deep learning', *Computers and Electronics in Agriculture*, Vol. 172, pp.105–339.
- Cong, Y. (2019) 'Novel event analysis for human-machine collaborative underwater exploration', *Pattern Recognition*, Vol. 96, pp.106–967.
- Cotter, E. and Staines, G. (2023) 'Observing fish interactions with marine energy turbines using acoustic cameras', *Fish and Fisheries*, Vol. 24, No. 6, pp.1020–1033.
- Dawson, H.L., Dubrule, O. and John, C.M. (2023) 'Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification', *Computers and Geosciences*, Vol. 171, pp.105284.
- Diamant, R. (2019) 'An active acoustic track-before-detect approach for finding underwater mobile targets', *IEEE Journal of Selected Topics in Signal Processing*, Vol. 13, No. 1, pp.104–119.
- Dong, J., Shangguan, X., Zhou, K., Gan, Y., Fan, H. and Chen, L. (2023) 'A detection-regression-based framework for fish keypoints detection', *Intelligent Marine Technology and Systems*, Vol. 1, No. 1, p.9.
- Fu, C., Liu, R., Fan, X., Chen, P., Fu, H., Yuan, W., Zhu, M. and Luo, Z. (2023) 'Rethinking general underwater object detection: datasets, challenges, and solutions', *Neurocomputing*, Vol. 517, pp.243–256.
- Gilbey, J. (2021) 'Life in a drop: sampling environmental DNA for marine fishery management and ecosystem monitoring', *Marine Policy*, Vol. 124, pp.104–331.

- Haider, A., Arsalan, M., Nam, S.H., Sultan, H. and Park, K.R. (2023) 'Computer-aided fish assessment in an underwater marine environment using parallel and progressive spatial information fusion', *Journal of King Saud University-Computer and Information Sciences*, Vol. 35, No. 3, pp.211–226.
- Hamzaoui, M., Ould-Elhassen Aoueilyne, M., Romdhani, L. and Bouallegue, R. (2023) 'An improved deep learning model for underwater species recognition in aquaculture', *Fishes*, Vol. 8, No. 10, p.514.
- Heidenreich, J.N., Bonatti, C. and Mohr, D. (2024) 'Transfer learning of recurrent neural network-based plasticity models', *International Journal for Numerical Methods in Engineering*, Vol. 125, No. 1, p.e7357.
- Huang, H., Zhou, H., Yang, X., Zhang, L., Qi, L. and Zang, A.Y. (2019) 'Faster R-CNN for marine organisms detection and recognition using data augmentation', *Neurocomputing*, Vol. 337, pp.372–384.
- Irfan, M., Jiangbin, Z., Iqbal, M. and Arif, M.H. (2021) 'A novel lifelong learning model based on cross-domain knowledge extraction and transfer to classify underwater images', *Information Sciences*, Vol. 552, pp.80–101.
- Jalal, A., Salman, A., Mian, A., Shortis, M. and Shafait, F. (2020) 'Fish detection and species classification in underwater environments using deep learning with temporal information', *Ecological Informatics*, Vol. 57, No. 1, p.101088.
- Knausgård, K.M., Wiklund, A., Sjørdalen, T.K., Halvorsen, K.T., Kleiven, A.R., Jiao, L. and Goodwin, M. (2022) 'Temperate fish detection and classification: a deep learning based approach', *Applied Intelligence*, Vol. 52, Nos. 1–14, pp.6988–7001.
- Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S. and Tao, D. (2019) 'An underwater image enhancement benchmark dataset and beyond', *IEEE Transactions on Image Processing*, Vol. 29, pp.4376–4389.
- Li, H., Yu, H., Zhang, P., Gao, H., Wei, S., Wei, Y., Xu, J., Cheng, S. and Wu, J. (2023) 'Detection of tiger puffer using improved YOLOv5 with prior knowledge fusion', *Information Processing in Agriculture*, Vol. 1, No. 1, p.1.
- Liu, R. (2020) 'Real-world underwater enhancement: challenges, benchmarks, and solutions under natural light', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 30, No. 12, pp.4861–4875.
- Moniruzzaman, M., Islam, S.M.S., Lavery, P. and Bennamoun, M. (2019) 'Faster R-CNN based deep learning for seagrass detection from underwater digital images', *2019 Digital Image Computing: Techniques and Applications (DICTA)*, IEEE, Perth, WA, Australia, pp.1–7.
- Paraschiv, M., Padrino, R., Casari, P., Bigal, E., Scheinin, A., Tchernov, D. and Fernández Anta, A. (2022) 'Classification of underwater fish images and videos via very small convolutional neural networks', *Journal of Marine Science and Engineering*, Vol. 10, No. 6, p.736.
- Qin, H. (2020) 'Autonomous environment and target perception of underwater offshore vehicles', *Fundamental Design and Automation Technologies in Offshore Robotics*, Academic Press, College of Shipbuilding Engineering, Harbin Engineering University, Harbin, China, pp.91–113.
- Rauf, H.T., Lali, M.I.U., Zahoor, S., Shah, S.Z.H., Rehman, A.U. and Bukhari, S.A.C. (2019) 'Visual features based automated identification of fish species using deep convolutional neural networks', *Computers and Electronics in Agriculture*, Vol. 167, p.105075.

- Song, W., Wang, Y., Huang, D., Liotta, A. and Perra, C. (2020) 'Enhancement of underwater images with a statistical model of background light and optimization of transmission map', *IEEE Transactions on Broadcasting*, Vol. 66, No. 1, pp.153–169.
- Stavelin, H., Rasheed, A., San, O. and Hestnes, A.J. (2021) 'Applying object detection to marine data and exploring the explainability of a fully convolutional neural network using principal component analysis', *Ecological Informatics*, Vol. 62, pp.101269.
- Sung, M. (2020) 'Realistic sonar image simulation using deep learning for underwater object detection', *International Journal of Control, Automation And Systems*, Vol. 18, No. 3, pp.523–534.
- Tanveer, A. and Ahmad, S.M. (2022) 'High fidelity modelling and GA optimized control of yaw dynamics of a custom built remotely operated unmanned underwater vehicle', *Ocean Engineering*, Vol. 266, p.112836.
- Tasyurek, M. and Arslan, R.S. (2023) 'RT-droid: A novel approach for real-time android application analysis with transfer learning-based CNN models', *Journal of Real-Time Image Processing*, Vol. 20, No. 3, pp.1–17.
- Vasile, G., Girard, A., d'Urso, G., De Oliveira, E., Hieramente, F. and Philippe, O. (2020) 'Active acoustic signature extraction by machine learning data analysis for underwater fish counters', *Global Oceans 2020: Singapore-US Gulf Coast*, October, IEEE, Biloxi, MS, USA, pp.1–4.
- Wan, S. and Goudos, S. (2020) 'Faster R-CNN for multi-class fruit detection using a robotic vision system', *Computer Networks*, Vol. 168, pp.107–036.
- Wang, X. (2019) 'Underwater object recognition based on deep encoding-decoding network', *Journal of the Ocean University of China*, Vol. 18, No. 2, pp.376–382.
- Wang, X., Xue, G., Huang, S. and Liu, Y. (2023) 'Underwater object detection algorithm based on adding channel and spatial fusion attention mechanism', *Journal of Marine Science and Engineering*, Vol. 11, No. 6, p.1116.
- Wang, Y., Yu, X., An, D. and Wei, Y. (2021) 'Underwater image enhancement and marine snow removal for fishery based on integrated dual-channel neural network', *Computers and Electronics in Agriculture*, Vol. 186, pp.106–182.
- Yang, M., Hu, K., Du, Y., Wei, Z., Sheng, Z. and Hu, J. (2020) 'Underwater image enhancement based on conditional generative adversarial network', *Signal Processing: Image Communication*, Vol. 81, pp.115–723.
- Yeh, C.H., Lin, C.H., Kang, L.W., Huang, C.H., Lin, M.H., Chang, C.Y. and Wang, C.C. (2021) 'Lightweight deep neural network for joint learning of underwater object detection and color conversion', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 11, pp.6129–6143.
- Yu, J., Wu, Z., Yang, X., Yang, Y. and Zhang, P. (2020) 'Underwater target tracking control of an untethered robotic fish with a camera stabilizer', *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 51, No. 10, pp.6523–6534.
- Yuan, J., Hu, Y., Sun, Y. and Yin, B. (2023) 'A multi-scale feature representation and interaction network for underwater object detection', *IET Computer Vision*, Vol. 17, No. 3, pp.265–281.
- Zhang, X. (2021) 'A marine organism detection framework based on the joint optimization of image enhancement and object detection', *Sensors*, Vol. 21, No. 21, p.7205.

- Zhao, Z., Liu, Y., Sun, X., Liu, J., Yang, X. and Zhou, C. (2021) ‘Composited FishNet: fish detection and species recognition from low-quality underwater videos’, *IEEE Transactions on Image Processing*, Vol. 30, pp.4719–4734.
- Zhou, J. (2022) ‘Underwater image enhancement method with light scattering characteristics’, *Computers and Electrical Engineering*, Vol. 100, pp.107–898.