Indian Journal of Geo Marine Sciences
Vol. 54 (09), September 2025, pp. 452-464
DOI: 10.56042/ijms.v54i09.18358

Research Article

Shrimp vision: Improved shrimp detection using underwater image enhancement and YOLOv8

M Sharma[a], V K Yadav*[,a], G K Birajdar[b] & S S Chaudhari[b]

[a]Department of Economics, Extension and Statistics, Central Institute of Fisheries Education, Mumbai, Maharashtra – 400 061, India
[b]Department of Computer Science and Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, Maharashtra – 400 706, India
*[E-mail: vinodkumar@cife.edu.in]

Received 29 March 2025; revised 16 August 2025

The advancement of automated approaches for detecting Pacific white shrimps in underwater environments is crucial for sustainable aquaculture practices. This study focuses on developing a novel methodology to enhance underwater shrimp images and detect shrimps using the YOLOv8 architecture. The specific objectives include developing a contrast-enhancement system to acquire clear underwater shrimp images and implementing the YOLOv8 model for shrimp detection. YOLOv8's superior accuracy, speed, and memory efficiency allow non-intrusive detection of bottom-dwelling shrimps using computer vision and machine learning without disrupting their natural habitat. However, underwater images suffer from poor contrast and limited visibility due to light attenuation and scattering, making object detection and recognition challenging. Hence, a fusion-based underwater image contrast enhancement algorithm is employed to improve the shrimp detection accuracy. The enhanced images obtained through the contrast-enhancement system are utilised for length and weight estimation of the shrimps using the YOLOv8 model. The results of the study demonstrate promising outcomes, with precision of 71.7 %, recall of 86.8 %, and mAP 50 scores of 87.8 % in shrimp detection. This comprehensive approach not only enhances the accuracy of shrimp detection in farm ponds but also contributes to the overall health monitoring and management of shrimp populations. The methodology employed in this study combines image enhancement techniques with state-of-the-art deep learning models to improve the efficiency and precision of shrimp detection in underwater environments.

[**Keywords**: Contrast enhancement system, Deep learning, Shrimp detection, Underwater image enhancement, YOLOv8]

Introduction

India ranks among the top five seafood-exporting nations, with exports reaching 1.36 million metric tons valued at $7.76 billion in 2021-22[ref. 1]. Frozen shrimp constitutes the most significant portion of these exports, accounting for over 53.18 % of the total quantity and 75.11 % of the total export value. The increasing shrimp cultivation in coastal regions necessitates continuous monitoring of environmental conditions to ensure optimal growth and survival. Computer vision and machine learning techniques have emerged as promising tools for monitoring aquatic species, yet underwater imaging remains a significant challenge due to light attenuation, scattering, and colour distortions[2,3]. These limitations result in low-contrast images with reduced visibility, making image enhancement and restoration critical[4,5].

Several techniques have been proposed to enhance underwater imagery, focusing on improving contrast, brightness, and colour balance. Enhancement methods such as histogram equalisation, brightness correction, and morphological filtering[6] are used to enhance visual clarity. In contrast, restoration techniques like the Dark Channel Prior (DCP)[7] and deep learning-based approaches reconstruct degraded images[8]. Methods such as WaterGAN[9] and UGAN[10] generate clear underwater images using deep learning; however, their computational complexity and limited adaptability across diverse underwater environments remain concerns.

Despite these advancements, a critical gap remains. Most existing underwater enhancement and detection models are tested under controlled laboratory or simulated environments and do not perform consistently in real-world aquaculture settings, particularly under turbid water conditions. Underwater images often suffer from colour distortion, low contrast, and poor clarity caused by selective light attenuation, absorption, and scattering. These issues hinder the accurate extraction of features necessary for shrimp monitoring. Several studies have attempted to mitigate these challenges through image enhancement techniques. Ghani *et al*.[11] introduced a method integrating a colour model with the Rayleigh distribution to minimise over- and under-enhancement regions, though the approach introduced noise into the output. Zhuang *et al*.[12] developed a method based on the retinex framework, combining variational processing, colour correction, and post-processing to improve the clarity of underwater images. While these methods address some issues, they often fail to generalise across varying environmental conditions.

Enhancement-based techniques also struggle to account for the complex physical parameters of the underwater environment, leading to inconsistencies in image quality. Restoration-based methods, including the Underwater Dark Channel Prior (UDCP) and red channel correction[7,8], rely on assumptions about light absorption that do not always hold true in turbid or bio-fouled environments. Additionally, many of these models have not been validated in live culture systems, limiting their applicability in real-time shrimp monitoring scenarios.

Recent advancements in deep learning have led to promising solutions for underwater image enhancement. Li *et al*.[9] introduced Water GAN, a Generative Adversarial Network (GAN) that synthesises underwater images with depth information to improve restoration. Li *et al*.[13] later proposed a weak underwater image colour-correction model using a multiterm loss function and a Cycle GAN to refine image quality. Fabbri *et al*.[10] developed the underwater GAN (UGAN), leveraging paired training data and a pix2pix-based model to enhance images. Despite their potential, these models often require substantial computational resources and still face challenges in adapting to diverse aquaculture conditions. Efficient training strategies such as spectral normalisation[14,15] offer promise but require further refinement for deployment in operational shrimp farms.

In parallel, detecting and tracking shrimp in aquaculture settings pose additional challenges due to environmental factors such as water turbidity, occlusions, and the presence of biofouling organisms, aerator bubbles, and overlapping marine elements[16,17]. Moreover, manual annotation of underwater images is prone to errors, leading to inconsistencies in training data[18]. Inconsistencies are often unavoidable due to the subjective nature of human annotations, though the impact can be mitigated by involving multiple experts and adopting majority voting schemes. However, manual annotation remains expensive and time-consuming, emphasising the need for automated detection methods.

Among object detection models, YOLO (You Only Look Once) has emerged as a leading real-time framework. Unlike Fast Region-based Convolutional Neural Network (R-CNN)[19], YOLO employs a unified detection framework that simultaneously predicts object class probabilities and bounding box coordinates, enhancing efficiency. YOLOv8, the latest iteration, demonstrates superior detection accuracy, inference speed, and memory efficiency compared to alternative models, such as Faster R-CNN, Cascade R-CNN, Single Shot MultiBox Detector (SSD), and RetinaNet[20]. YOLOv8 also integrates improved feature extraction and multi-scale prediction capabilities, making it well-suited for complex underwater environments.

While prior studies have applied deep learning models to fish detection in controlled settings, including YOLOv2 for fish classification[21] and SFCNet for shrimp fry counting[22], there remains a lack of research targeting real-world shrimp culture environments. For instance, Roy *et al*.[23] achieved high detection accuracy using YOLOv8 with MaxRGB filtering; however, the study was conducted under constrained experimental conditions. The absence of integrated solutions that combine underwater image enhancement with real-time shrimp detection in actual, turbid culture tanks highlights a major research gap.

This study aims to fill this gap by developing an automated system for detecting Pacific white shrimp (*Litopenaeus vannamei*) under real aquaculture conditions. Given the bottom-dwelling behaviour of shrimp, the system employs a non-intrusive imaging approach to avoid disturbing their normal physiological activity. The specific objectives are to enhance underwater images, obtain clearer visuals of shrimp in turbid environments, and accurately detect them using a real-time deep learning model (YOLOv8). This work seeks to lay the foundation for practical, automated monitoring systems that support shrimp health and growth management in operational aquaculture farms.

Methodology

The proposed approach is based on the Ancuti *et al*.[24] model. This system is simple, computationally economical, and capable of achieving comparatively quick results on standard hardware. Underwater imagery degradation[25] follows a combined multiplicative and additive model; thus, traditional image enhancement techniques such as colour correction, white balance adjustment, and histogram equalisation are inherently limited in their effectiveness when applied to underwater image restoration. An alternative to directly filtering the input image is a developed fusion-based technique inspired by the inherent properties of the original image (represented by the weight maps). Figure 1 shows the outline of the proposed shrimp detection approach.

Initially, the damaged image is white-balanced to remove colour casts and bring back the original appearance of the underwater image. Subtracting some of the unwanted noise greatly improves this version, which has been partially restored. This filtered version serves as the source for the second input, which renders the details across the whole intensity range. Multiple weight maps serve as the engine for the developed fusion-based augmentation technique. The algorithm's weight maps evaluate several picture attributes that define the spatial pixel associations. These weights give pixels more values to accurately portray the desired image characteristics. Lastly, this method is resilient to artefacts because it is multi-resolution in nature and can handle dynamic scenes.

Image acquisition

The shrimp culture underwater images were captured with a GoPro Hero 10. This camera was selected due to its combination of waterproofing, anti-shake technology, and strong low-light performance. The image acquisition process contains three steps:

a. Video capturing: The videos were taken daily for 20 days. The resolution of the videos was 5.3 K. The videos captured by the GoPro camera were in H.265 (High Efficiency Video Coding) format.

b. Conversion: The H.26 videos were converted to MP4 videos by using the Video- Proc app on the Google platform. The conversion was required as the image cannot be directly extracted from HEVC formats.

c. Extraction of image: The images from the videos were extracted using the same video Proc app. The images were extracted at 60 frames/sec. 1,19000 images have been extracted, and a dataset has been made classifying the different turbidity levels of the saline water. A total of 856 images were taken for enhancement and restoration; these images represented different turbidity levels.

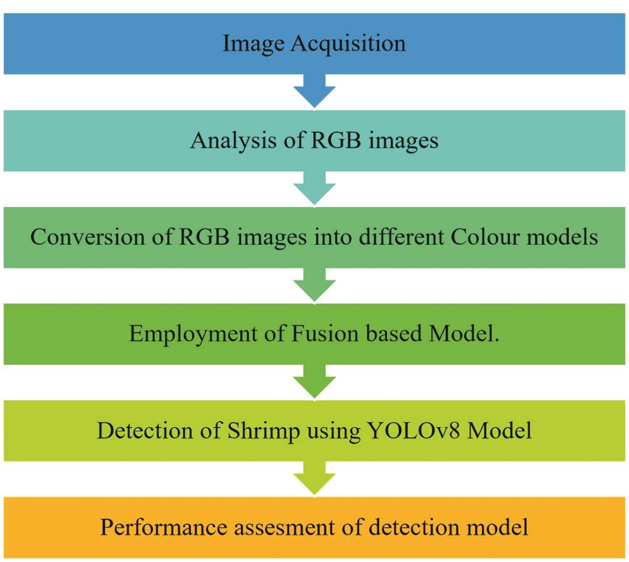

Fig. 1 — Workflow of shrimp detection approach used in current study

Image processing

I. Analysis of RGB images: Analysis of RGB (Red, Green, Blue) images provided valuable insights into various aspects of visual content, including colour distribution, intensity variations, and structural details. By examining the individual channels of an RGB image, important information about the composition and characteristics of the scene captured in the underwater images was obtained. Firstly, analysing the colour distribution across the RGB channels revealed the dominant hues in the image. Different scenes may exhibit distinct colour signatures, with certain colours being more prevalent than others. Figure 2 illustrates the input shrimp image, its corresponding RGB components and the histogram.

Thus, by comparing histograms of different colour channels, the colour imbalances or inconsistencies were identified that needed correction. For instance, the absorption and scattering of light in water often cause underwater photos to appear green and blue. By studying the intensity levels in each channel, it was observed that the red component contributed more to the image than the blue and green components. Additionally, the colour channel analysis provides the foundation for more advanced image processing techniques, such as channel mixing, selective colour adjustments, and targeted enhancements to specific colour components.

II. Conversion of RGB images into different colour models: Converting RGB colour model images to

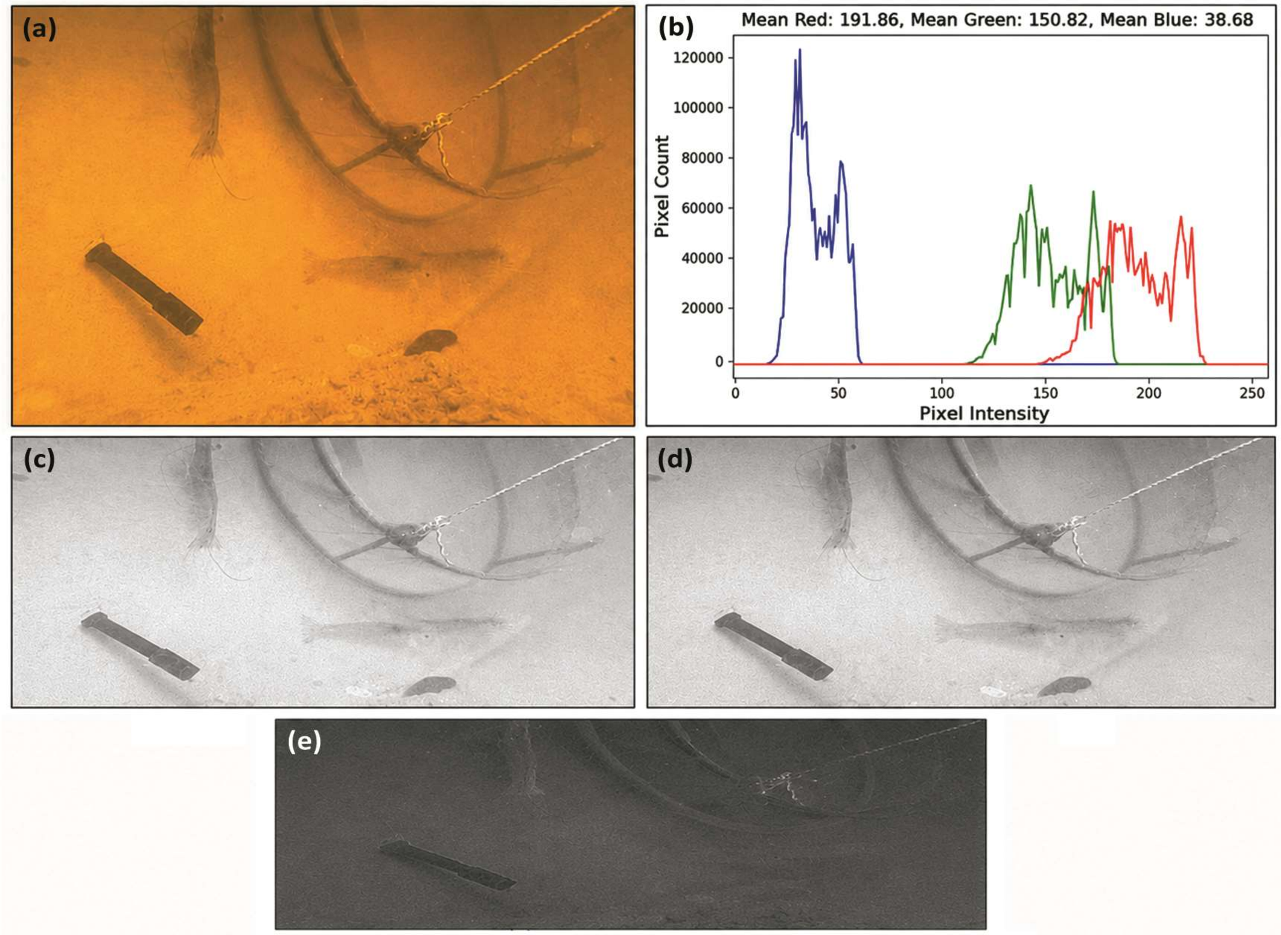

Fig. 2 — Shrimp image and RGB components: a) Underwater shrimp image, b) RGB histogram, c) Red component, d) Green component, and e) Blue component

different colour models offers advantages such as perceptual uniformity, efficient colour separation, and specialised analysis. Models like CIELAB provide consistency in colour perception, while HSV separates hue, saturation, and value for intuitive adjustments. Additionally, models like YCbCr enable efficient compression and transmission of colour information. These conversions led understanding of the specific application requirements, optimised image processing tasks, and enhanced overall versatility in image analysis and manipulation.

III. Employment of fusion-based model: Proposed method focuses on improving a single degraded image by using a fusion technique. The main idea behind image fusion is to merge multiple images in a way that brings out their most important features, resulting in a final image that showcases the best aspects of each input. In this approach, two different versions of the original degraded image were created to use as inputs for the fusion process.

Rather than relying on complex techniques that require detailed physical models of the scene, a simpler, more practical approach that is easy to apply in many situations is used in this study. The first version is an image with adjusted colours to make them more accurate. The second version enhances the image by boosting contrast and reducing noise. Using these two inputs, the overall quality of the image can be improved significantly, ensuring that the final result looks as good as possible in every part.

One essential pre-processing step that addresses the common colour distortion in underwater images, which is mostly caused by light absorption and scattering, is red colour adjustment. The objective of this step is to restore colour balance and integrity by selectively altering the red channel's intensity.

$$IEnh(i, j) = Ir(i, j) + \alpha\,(gmean - rmean) \times (1 - Ir(i, j))Ig(i, j) \quad \ldots (1)$$

The red channel intensity at pixel *Ir (i, j)* and the red channel intensity within the same local window are represented by *r* mean. The average intensity value of the green channel within a local window centred on pixel *Ig (i, j)* is represented by *g* mean. The adjustment factor controlling the magnitude of compensation is represented by α. A key component of image processing is white balance, which improves the appearance of a picture by removing unwanted colour casts caused by different lighting conditions. However, in this study, the depth is generally 1.5 – 2 m and the effectiveness of white balancing diminishes notably due to challenges in restoring absorbed colours. Moreover, underwater environments are characterised by a substantial lack of contrast, attributable to the limited propagation of light within this medium due to turbidity and water salinity.

RGB to LAB colour space conversion is applied to the input image, which is represented in the RGB colour space, to convert it to the LAB colour space to improve colour perception and enable white balancing. Three channels make up the LAB colour space: L represents brightness, whereas A and B represent colour information. Then, inside the LAB colour space, mean intensity values are computed for each channel (L, A, and B), offering information on overall brightness and colour balance. The procedure is then set by determining the highest and lowest intensity values for each channel to normalise the intensity values across a shared scale and modify saturation. Additionally, a small percentage (0.5 %) of the intensity range is clipped at both ends to remove extreme values that could distort colour balance. Based on the ratio of the maximum to average intensity values, adjustment factors are computed for each channel. These adjustment factors scale the intensity values to achieve a balanced distribution across colour channels. Finally, the intensity values for each channel are adjusted based on the calculated factors, ensuring proper colour balance where white objects appear neutral without any colour casts. This comprehensive process ensures a more refined image appearance while maintaining original colour integrity.

Gamma correction, also known as gamma adjustment or gamma encoding, is an image processing technique that modifies an image's brightness and contrast by transforming pixel values nonlinearly. By modulating the gamma value, this technique effectively enhances image brightness and contrast, optimising perceptual quality and facilitating superior visualisation of underwater scenes. Afterwards, gamma correction is applied by raising each pixel's intensity value to the given gamma value. Additionally, the result is multiplied by the scaling factor alpha, which allows for further adjustment of the gamma-corrected image's intensity values.

Sharpening techniques play a pivotal role in enhancing image detail and edge delineation, particularly in underwater imagery characterised by reduced contrast and blurred features. By selectively enhancing high-frequency components, sharpening filters improve image sharpness and clarity, thereby enhancing overall visual acuity and interpretability. The input image is blurred using a 5*5 Gaussian kernel with a standard deviation of 3. The unsharp mask is generated by subtracting the blurred image from the original image. Histogram stretching is used to boost the contrast of the unsharp mask, enhancing the image's details. The sharpened image is obtained by blending the histogram-stretched unsharp mask with the original image.

The blending process averages the pixel values from two images to create the final image. The following steps are performed to obtain different weights:

- Weights are computed before image fusion to establish each input image's contribution at various spatial scales.
- Different weights are used to achieve the optimal outcome based on factors such as Laplacian contrast, saliency, and saturation.
- Laplacian contrast measures the sharpness of image details, saliency identifies visually significant regions, and saturation indicates the intensity of colours.
- The Laplacian contrast weight (WL) is determined by first applying a Laplacian filter to each luminance channel and then calculating the absolute value of the filter's output, thereby enhancing global contrast.
- The Local Contrast Weight (WLC) measures the correlation between each pixel and the average luminance of its neighbourhood, boosting local contrast.
- WLC is computed using the standard deviation between a pixel's luminance and the average luminance of its neighbouring pixels.

The Saliency Weight (WS) accentuates discernible objects that may otherwise become less visible in an underwater scene. The biological notion of centre-surround contrast serves as the basis for this method, which is renowned for its computational efficiency. But it favours highlighted parts more often than other parts, which could lead to errors. This study presents the exposedness map to improve the accuracy of the results by protecting mid-tones that might be harmed under specific conditions. This additional measure ensures a more comprehensive treatment of the scene's salient features while maintaining fidelity to the original image.

The conventional multi-scale Laplacian pyramid decomposition is employed in the proposed study. Using the Laplacian operator, each input image is represented by this linear decomposition technique as the sum of patterns calculated at different scales. To create low-pass-filtered versions of the original images, the inputs are first convolved with a Gaussian kernel. The Gaussian kernel's standard deviation is gradually raised to adjust the cutoff frequency. After low-pass filtering of the original image, the difference between the filtered and original images is computed to obtain the pyramid levels. For neighbouring Gaussian pyramid levels, this procedure is literally performed. The outcome yields quasi-bandpass copies of the image in the form of a Laplacian pyramid. By applying the Laplacian operator at multiple scales, current method decomposes each input into a pyramid. A Gaussian pyramid is also calculated for every normalised weight map. As the Gaussian and Laplacian pyramids have the same number of levels, the fusion of the Laplacian inputs and Gaussian-normalised weights occurs at each level separately, resulting in the fused pyramid. This process ensures effective fusion of information across multiple scales while preserving important image characteristics.

Detection of shrimp using YOLOv8 model

The YOLO series employs the entire image as input and immediately regresses the bounding box's position and category in the output layer, in contrast to candidate region-based target identification models, for example, Faster-RCNN by Ren *et al.*[27]. By using this technique, the two-stage target detection and classification challenge is reduced to a single-stage regression issue, enabling the model to finish end-to-end processing inside the task. The detailed YOLOv8 architecture is outlined in Figure 3. While YOLOv8's core remains comparable to that of YOLOv5, various changes have been made, most notably in the C2f module (formerly known as the CSP Layer). By fusing contextual data with high-level features, the C2f module, a cross-stage partial barrier with two convolutions, improves detection accuracy.

This arrangement enhances the model's overall accuracy by enabling each branch to focus on its designated task. To describe the likelihood that an object would be included in the bounding boxes, the sigmoid function was used as the activation function for the objectness score in the YOLOv8 output layer. The softmax function is used to illustrate the likelihood of things falling into every possible category YOLOv8 uses a decoupled head and an anchor-free model to process Object-detection, classification, and regression tasks independently.

YOLOv8 uses binary cross-entropy for the classification loss and CIoU and DFL loss functions for the bounding-box loss[27,28]. Especially when it comes to small objects, these losses have improved object detection performance. YOLOv8 also provides a semantic segmentation model known as the YOLOv8-Seg model. Unlike the conventional YOLO neck architecture, the C2f module comes after the backbone, CSPDarknet53, which serves as a feature extractor. Two segmentation heads succeed in the C2f module by learning to anticipate the semantic segmentation masks for the supplied image. With five detection modules, a prediction layer, and detection heads similar to YOLOv8, the model is quite similar.

Performance assessment parameters

Performance indicators are crucial for evaluating the efficacy of the target detection algorithm. Prediction accuracy is indicated by the precision rate, which calculates the likelihood of a positive sample among expected positive samples. Recall rate indicates overall prediction accuracy by calculating the probability of correctly classifying a positive sample as positive.

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

This formula defines *TP* (True Positives) as a correctly identified shrimp, *FN* (False Negatives) as a sample misidentified as a background, *TN* (True Negatives) as a correctly identified background, and *FP* (False Positives) as a sample misidentified as a shrimp. The Average Precision (AP) is a commonly used performance statistic for target detection systems. Instances denote the total number of

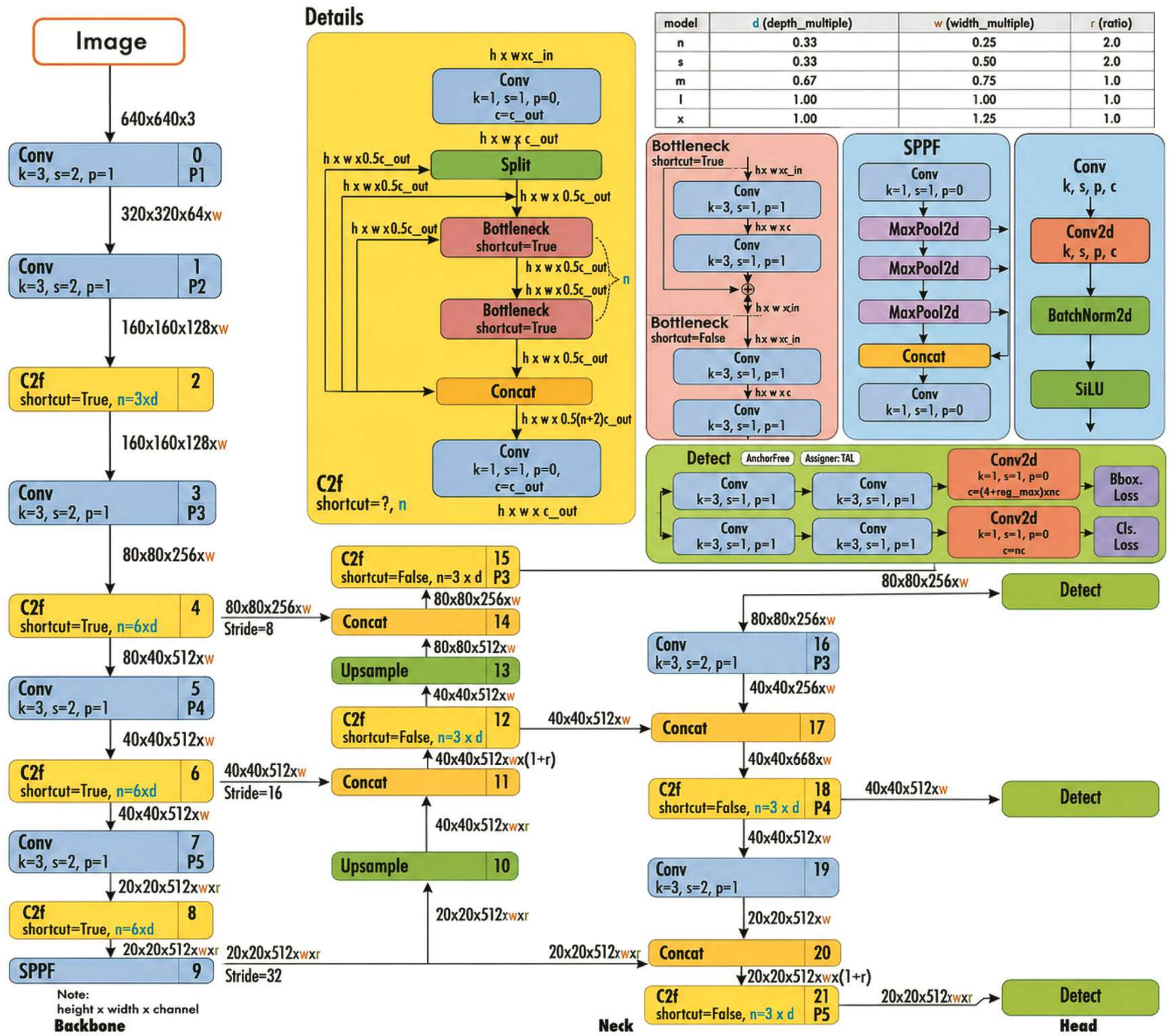

Fig. 3 — YOLOv8 architecture[29]

instances or occurrences of objects detected across all images. The percentage of correctly classified values, or accuracy, is a crucial parameter that indicates how often the shrimp is successfully detected. It is calculated by dividing the total value by the sum of all true values.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Mean Average Precision or mAP50 is a threshold of 0.5 for the Intersection over Union (IoU), which is used to calculate Mean Average Precision or mAP50. IoU is used to compare predicted bounding boxes for objects in an image with ground-truth bounding boxes in object detection. When computing mAP50, only detections with an IoU $\geq$ 0.5 are considered valid. The term "mAP50-95" refers to the mean Average Precision, which can be determined by averaging the AP scores across a range of IoU values from 0.5 to 0.95, with a 0.05 leap value. Compared to mAP50, this statistic takes into account a wider range of IoU thresholds. It provides a more thorough evaluation of how well the model performs at various crossovers between the ground truth and predicted bounding boxes.

Results

Image enhancement and visual inspection

The proposed fusion-based underwater image enhancement method produced visually superior results, significantly improving the global contrast,

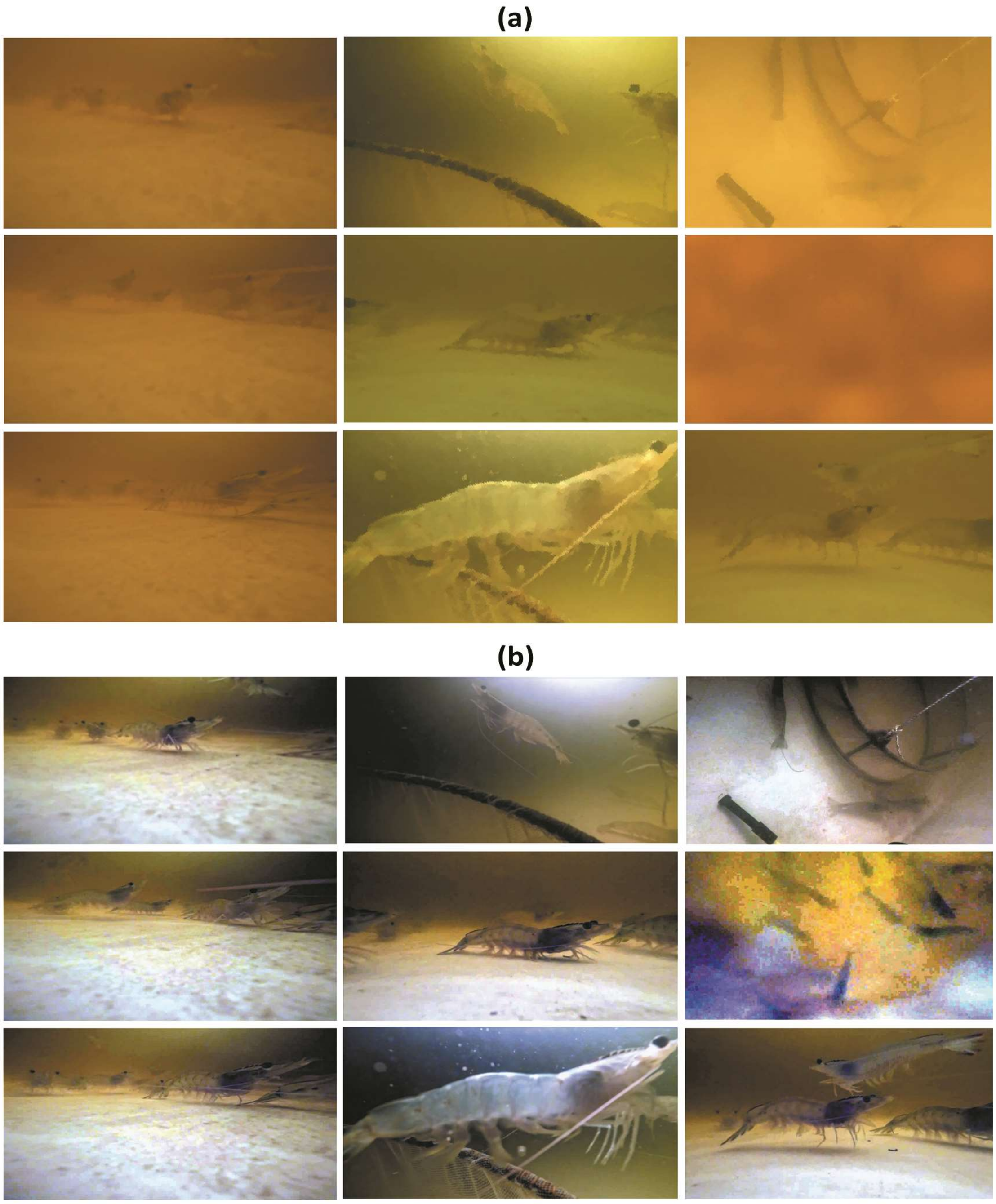

Fig. 4 — Fusion-based underwater image enhancement results: a) Original shrimp images; and b) Enhanced shrimp images

colour accuracy, and clarity of shrimp images. Upon visual inspection, the enhanced images showed a noticeable reduction in colour cast and the elimination of halos and colour distortion, common issues in underwater imagery.

Figure 4 illustrates the difference between the original and enhanced shrimp images, where the enhanced samples appear clearer and more distinguishable. However, the technique displayed limitations in extremely turbid conditions illuminated by natural light, where the restoration of distant scene areas proved unreliable due to insufficient illumination.

YOLOv8 model training and performance

The enhanced images were used to train the YOLOv8 network model, significantly boosting the model's performance.

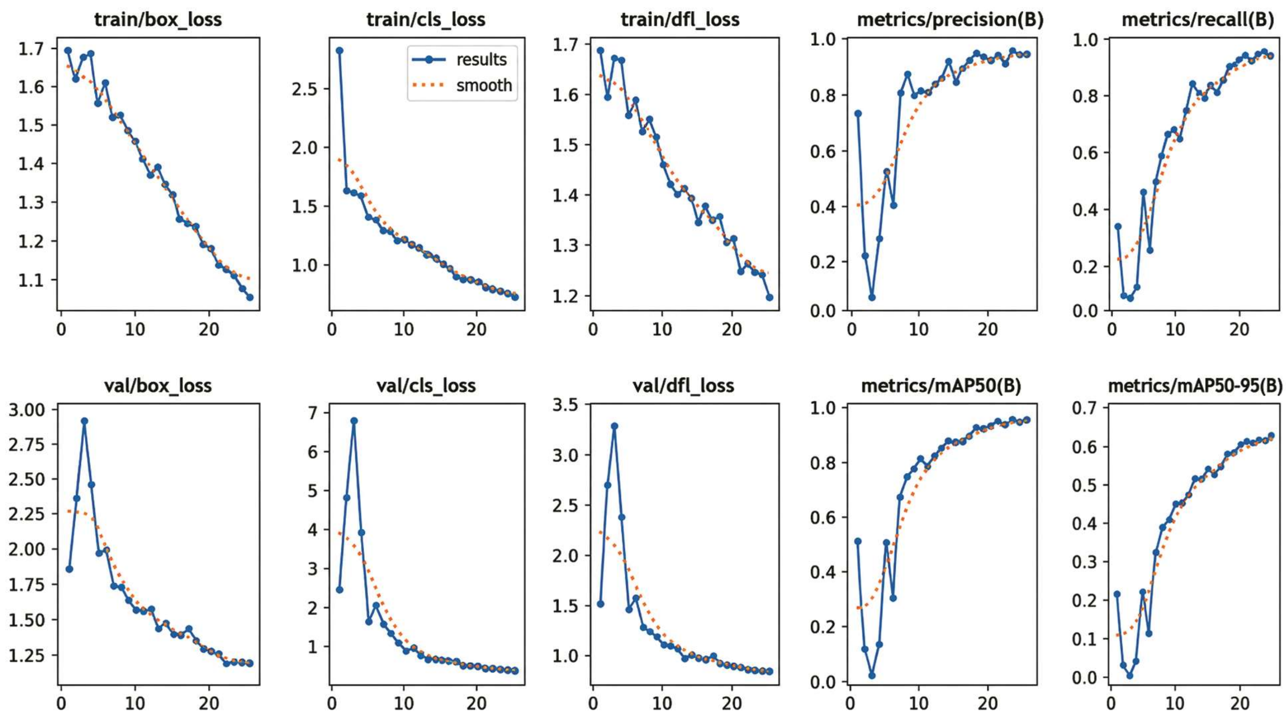

Fig. 5 — The loss graphs during the training and validation at 25 epochs

Figure 5 presents the loss curves during training and validation at 25 epochs, demonstrating a consistent downward trend, indicating effective learning. The precision and recall values gradually increased throughout the training, confirming that the model became more accurate and thorough in its predictions. The mAP scores also improved progressively, highlighting the model's growing ability to detect objects with greater reliability. Overall, the YOLOv8 architecture showed steady performance gains, validating the effectiveness of the enhanced images in improving shrimp detection accuracy in underwater conditions.

For shrimp detection, a custom dataset comprising 856 images was created. These images were processed separately in their original and enhanced forms. During the annotation phase, it was observed that the original images enabled clear and accurate labelling in only 628 images, resulting in a 73 % annotation rate. In contrast, the enhanced images provided clearer visual information, enabling the accurate annotation of 754 images, corresponding to an 88 % annotation rate. This improvement underscores the impact of image enhancement facilitating better and more precise data labelling. The dataset was divided into a train-test-validation split of 70:20:10, with training carried out over 25 epochs, optimising the model's generalisation capability. The detection performance of YOLOv8 model is shown in Figure 6.

Detection performance metrics

The effectiveness of the image enhancement process on the shrimp detection performance was quantified through key metrics. Table 1 shows significant increases in precision, recall, and mAP values after enhancement. The precision improved from 64.3 % to 71.7 %, while recall increased notably from 73.2 % to 86.8 %, indicating the model's enhanced ability to correctly identify shrimp. The mAP 50 score, representing the model's overall accuracy, improved from 73.2 % to 87.8 %, demonstrating a substantial gain in detection reliability. Moreover, the mAP 50-95, which reflects the model's consistency across various Intersection over Union (IoU) thresholds, increased from 41.8 % to 58.5 %, highlighting the contribution of image enhancement to improved object detection across various conditions.

Impact of hyperparameter tuning

The influence of hyperparameter variations, such as the number of epochs, was also evaluated. The results are shown in Table 2. The results revealed that increasing the number of epochs led to significant

Fig. 6 — Detection performance of YOLOv8 model

Table 1 — Effect of underwater image enhancement on shrimp detection performance measures

Parameter	Before enhancement	After enhancement
Validation images	125	151
Instances	224	445
Precision	0.643	0.717
Recall	0.732	0.868
mAP 50	0.732	0.878
mAP50-95	0.418	0.585

Table 2 — Number of epochs and shrimp detection performance parameter variations

Parameter	10 epoch	100 epoch
Validation images	151	151
Instances	445	445
Precision	0.73	0.7[21]
Recall	0.843	0.917
mAP 50	0.838	0.886
mAP50-95	0.537	0.623

performance gains. At 100 epochs, recall improved by 7.4 %, rising from 84.3 % to 91.7 %, while the mAP 50 score increased by 4.8 %, reaching 88.6 %. However, a slight decline in precision was observed, from 73.0 % to 72.1 %. This minor drop in precision indicates that while the model became more sensitive (higher recall), it also produced slightly more false positives. These findings demonstrate that while increasing the number of epochs enhances the model's generalisation and recall capacity, it can marginally affect precision.

Influence of Dataset split ratios

The impact of varying the train-test-validation configurations on performance metrics was also evaluated and the results are shown in Table 3. Two dataset splits were compared: 70:15:15 and 80:10:10. The configuration with a larger training set (80:10:10) yielded superior results. The mAP 50 score increased

Table 3 — Different train-validation-test configurations and statistical measures

Parameter	(70:15:15)	(80:10:10)
Validation images	113	75
Instances	352	179
Precision	0.738	0.813
Recall	0.89	0.827
mAP 50	0.898	0.93
mAP50-95	0.595	0.636

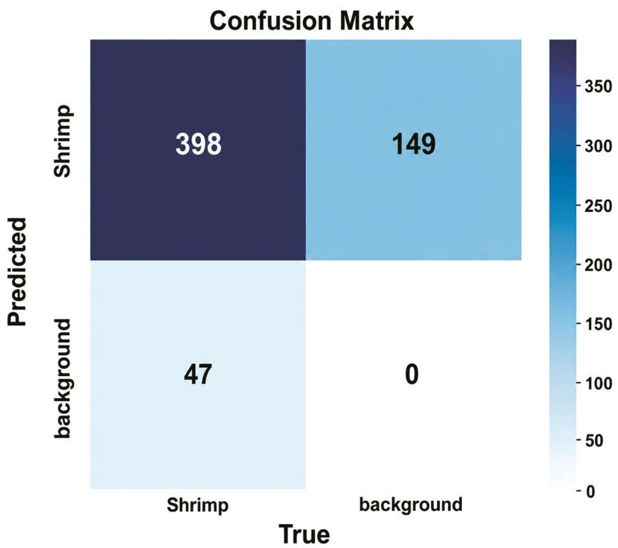

Fig. 7 — Confusion matrix of the YOLOv8 model

to 93.0 %, marking a 3.2 % gain over the 70:15:15 split and a 5.2 % improvement over the original 70:20:10 configuration. This indicates that a larger training set allows the model to learn more effectively, resulting in higher detection accuracy and better generalisation on unseen data. Precision and recall also improved with the 80:10:10 split, confirming the advantage of providing the model with more training data.

Confusion matrix analysis

The model's performance was further evaluated using a confusion matrix, as shown in Figure 7. The matrix showed that the model correctly identified shrimp in 369 instances, classified as True Positives (TP). However, it failed to detect shrimp in 149 instances, which were misclassified as background, categorised as False Negatives (FN). Additionally, 47 background instances were incorrectly predicted as shrimp, resulting in False Positives (FP). Interestingly, no True Negative (TN) instances were recorded, indicating the absence of correctly classified background samples, which is expected given the shrimp-focused nature of the dataset. The confusion matrix effectively highlights the model's strong shrimp detection capability but also reflects areas where further refinement could enhance accuracy by reducing false negatives and false positives.

Discussion

The results clearly demonstrate that the proposed fusion-based image enhancement significantly improved the quality and clarity of underwater shrimp images. The enhanced images facilitated more accurate manual annotation, increasing from 73 % for the original images to 88 % for the enhanced ones. This improved clarity translated into better training data, enabling the YOLOv8 model to achieve higher precision, recall, and mAP scores.

Although direct experimental comparisons with other underwater image enhancement methods were not performed in this study, findings are consistent with those reported in prior works such as Ancuti *et al.*[24], where fusion-based enhancement demonstrated superior contrast amplification, colour correction, and preservation of structural details over traditional techniques, including polarisation-based and single-image dehazing methods. The improvements observed both visually and in downstream model performance are consistent with the reported benefits, thereby reinforcing the efficacy of this approach for aquaculture-specific underwater imaging.

The YOLOv8 model demonstrated robust performance improvements when trained on enhanced images. The consistent rise in recall and mAP values, particularly at higher epochs, indicates the model's capacity to generalise effectively with more training iterations. However, the slight decline in precision at 100 epochs suggests that while the model became more sensitive to detecting shrimp, it also produced a few more false positives. The confusion matrix further confirmed the model's reliable classification performance, while highlighting areas where misclassifications (false positives and false negatives) still occurred.

The analysis of hyperparameter variations revealed that longer training with more epochs improved the model's recall and mAP scores but marginally reduced precision. This trade-off suggests that while longer training enhances the model's sensitivity, it may slightly compromise its specificity. The comparison of dataset split configurations demonstrated that a larger training set, as in the 80:10:10 configuration, yielded the best results. The higher mAP 50 score of 93.0 % in this configuration underscores the benefit of providing the model with more training data, allowing it to learn more complex features and generalise better.

This computational efficiency, combined with its ability to improve image clarity and support AI-driven shrimp detection tasks, underscores its practical suitability for aquaculture environments where real-time, low-cost solutions are preferred. Despite the promising results, the study has certain limitations. The custom dataset, while effective for training and evaluation, may not encompass the full range of environmental conditions encountered in real-world shrimp culture settings. Variations in water quality, lighting, and the presence of other aquatic species could potentially affect the model's performance in practical applications. Additionally, the manual annotation process, although thorough, introduces a degree of subjectivity, which may result in occasional inconsistencies in the training data. Moreover, while robust, the YOLOv8 model may face challenges in highly cluttered environments, where occlusion by other objects can reduce detection accuracy. The computational resource requirements of the deep learning model may also limit its adoption by smaller-scale aquaculture operations with limited infrastructure.

Conclusion

This study demonstrates significant advancements in underwater imaging and shrimp identification, specifically targeting Pacific white shrimp (*Litopenaeus vannamei*). By integrating fusion-based image enhancement with the YOLOv8 detection model, the system effectively improved underwater image quality, achieving 71.7 % precision, 86.8 % recall, and an impressive 87.8 % mAP 50 score. The enhanced images not only facilitated more accurate shrimp detection but also enabled better monitoring of their size and health, which is critical for aquaculture operations. The non-invasive nature of the method ensures that shrimp behaviour and habitat remain undisturbed, promoting sustainable aquaculture practices. Despite its effectiveness, the study acknowledges the need for a more diverse dataset and automated annotation techniques to enhance generalisability. Future research should explore transformer-based models and real-time monitoring systems for improved accuracy and operational efficiency. Overall, this work offers a robust solution for shrimp detection, thereby enhancing aquaculture management and productivity.

Acknowledgements

The authors thank ICAR-Central Institute of Fisheries Education for providing the facilities we needed to complete the study.

Conflict of Interest

Authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Statement

The authors declare that this research complies with all applicable ethical standards. The study did not involve human participants, human tissues, or experimental procedures on animals requiring ethical approval. Data acquisition was limited to non-invasive underwater imaging conducted under normal aquaculture management practices.

Author Contributions

MS: Lab work, data analysis, and drafting of the manuscript; GKB: Reviewing and editing of the manuscript; VKY & SSC: Provided guidance for the entire research study process and corrected the manuscript draft.

References

1 Fisheries Statistics Division, *Handbook on Fisheries Statistics, 2022*, (Department of Fisheries, Ministry of Fisheries, Animal Husbandry and Dairying, Government of India), 2022, pp. 218.

2 Raveendran S, Patil M D & Birajdar G K, Underwater image enhancement: a comprehensive review, recent trends, challenges and applications, *Artif Intell Rev*, 54 (7) (2021) 5413–5467. https://doi.org/10.1007/s10462-021-10025-z

3 Schettini R & Corchs S, Underwater image processing: State of the art of restoration and image enhancement methods, *EURASIP J Adv Signal Process*, 2010 (2010) 1-14, Art No 746052. https://doi.org/10.1155/2010/746052

4 Liu J G & Mason P J, Digital Image and Display, In: *Essential Image Processing and GIS for Remote Sensing*, (Wiley Interscience, USA), 2013, pp. 3–8. https://doi.org/10.1002/9781118687963.ch1

5 Raveendran S, Patil M D & Birajdar G, Comparative analysis of hand-crafted feature descriptors and deep learning architectures for underwater coral-reef classification, *Proc 2023 6th Int Conf on Advances in Science and Technology (ICAST)*, 2023, pp. 569–574. https://doi.org/10.1109/ICAST59062.2023.10454953

6 Acharya T & Ray A K, *Image Processing - Principles and Applications*, (Wiley Interscience, USA), 2005, pp. 105–130

7 He K, Sun J & Tang X, Single image haze removal using dark channel prior, *IEEE Trans Pattern Anal Mach Intell*, 33 (12) (2011) 2341–2353. https://doi.org/10.1109/TPAMI.2010.168

8 Galdran A, Pardo D, Picón A & Alvarez-Gila A, Automatic red-channel underwater image restoration, *J Vis Commun Image Represent*, 26 (2015) 132–145. https://doi.org/ 10.1016/j.jvcir.2014.11.006

9 Li J, Skinner K A, Eustice R M & Johnson-Roberson M, WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images, *IEEE Robot Autom Lett*, 3 (1) (2017) 387–394. https://doi.org/10.1109/LRA.2017.2730363

10 Fabbri C, Islam M J & Sattar J, Enhancing underwater imagery using generative adversarial networks, *Proc 2018 IEEE Int Conf on Robotics and Automation (ICRA)*, 2018, pp. 7159–7165.

11 Abdul Ghani A S & Mat Isa N A, Underwater image quality enhancement through integrated color model with Rayleigh distribution, *Appl Soft Comput*, 27 (2015) 219–230. https://doi.org/10.1016/j.asoc.2014.11.020

12 Fu X, Zhuang P, Huang Y, Liao Y, Zhang X-P, *et al.*, A retinex-based enhancing approach for single underwater image, *Proc 2014 IEEE Int Conf on Image Processing (ICIP)*, 2014, pp. 4572–4576. https://doi.org/ 10.1109/ ICIP.2014.7025927

13 Li C, Guo J & Guo C, Emerging from water: Underwater image color correction based on weakly supervised color transfer, *IEEE Signal Process Lett*, 25 (3) (2018) 323–327. https://doi.org/10.1109/LSP.2018.2792050

14 Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V & Courville A, Improved training of Wasserstein GANs, *Proc 31st Int Conf on Neural Information Processing Systems (NIPS'17)*, (Curran Associates Inc., Red Hook, NY, USA), 2017, pp. 5769–5779.

15 Miyato T, Kataoka T, Koyama M & Yoshida Y, Spectral Normalization for Generative Adversarial Networks, *arXiv preprint,* 2018. https://arxiv.org/abs/1802.05957

16 Marini S, Fanelli E, Sbragaglia V, Azurro E, Del Rio J, *et al.*, Tracking fish abundance by underwater image recognition, *Sci Rep*, 8 (2018) 1–12, Art No 13748. https://doi.org/10.1038/s41598-018-32089-8

17 Abe S, Takagi T, Torisawa S, Abe K, Habe H, *et al.*, Development of fish spatio-temporal identifying technology using segnet in aquaculture net cages, *Aquacult Eng*, 93 (2021) p. 102146. https://doi.org/10.1016/ j.aquaeng.2021.102146

18 Atienza-Vanacloig V, Andreu-García G, López-García F, Valiente-González J M & Puig-Pons V, Vision-based discrimination of tuna individuals in grow-out cages through a fish bending model, *Comput Electron Agric*, 130 (2016) 142–150. https://doi.org/10.1016/j.compag.2016.10.009

19 IEEE, *Proceedings of IEEE International Conference on Computer Vision*, 20-23 June 1995, Cambridge, MA, USA, 1995. https://doi.org/10.1109/ICCV.1995.466933

20 Zariful M, Tashim N, Lim T, Basri A, Fakhrurroja H, *et al.*, Analysis of deep learning algorithms for prawn aquaculture in a challenging environment, *Proc 6th Int Conf on Applied Computational Intelligence in Information Systems (ACIIS 2023)*, 2023, pp. 133–137

21 Park J-H & Kang C, A study on enhancement of fish recognition using cumulative mean of YOLO network in underwater video images, *J Mar Sci Eng*, 8 (2020) 1–15. https://doi.org/10.3390/jmse8110952

22 Peng X, Zhou T, Zhang Y & Zhao X, Automatic shrimp fry counting method using multi-scale attention fusion, *Sensors*, 24 (2024) 1-24. https://doi.org/10.3390/s24092916

23 Dakhil R & Khayeat A, Review on deep learning techniques for underwater object detection, *Proc 2022 Conf on Computer Science and Information Technology (CSIT)*, 2022, pp. 49–63. https://doi.org/10.5121/ csit.2022.121505

24 Ancuti C, Ancuti C O, Haber T & Bekaert P, Enhancing underwater images and videos by fusion, *Proc 2012 IEEE Conf on Computer Vision and Pattern Recognition*, 2012, pp. 81–88.

25 Schettini R & Corchs S, Underwater image processing: State of the art of restoration and image enhancement methods, *EURASIP J Adv Signal Process*, 2010 (2010) 1-14, Art no 746052. https://doi.org/10.1155/2010/746052

26 Ren S, He K, Girshick R & Sun J, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Trans Pattern Anal Mach Intell*, 39 (6) (2017) 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

27 Zheng Z, Wang P, Liu W, Li J, Ye R, *et al.*, Distance IoU loss: Faster and better learning for bounding box regression, *Proc AAAI Conf Artif Intell*, 2020, pp. 12993–13000.

28 Li X, Wang W, Wu L, Chen S, Hu X, *et al.*, Generalized Focal Loss: Learning qualified and distributed bounding boxes for dense object detection, *NIPS'20: Proc 34th Int Conf Neural Inf Process Syst,* 2020, pp. 21002-21012

29 Sattar K, Maqsood U, Hussain Q, Majeed S, Kaleem S, *et al.*, Soil texture analysis using controlled image processing, *Smart Agr Technol*, 2024, pp. 1–15.