A novel automated approach for fish biomass estimation in turbid environments through deep learning, object detection, and regression

Article *in* Ecological Informatics · June 2024

DOI: 10.1016/j.ecoinf.2024.102663

CITATIONS
6

READS
133

5 authors, including:

Jansi Rani Sella Veluswami
Sri Sivasubramaniya Nadar College of Engineering
24 PUBLICATIONS **128** CITATIONS
SEE PROFILE

Iacovos Ioannou
University of Cyprus
67 PUBLICATIONS **259** CITATIONS
SEE PROFILE

Vasos Vassiliou
University of Cyprus
187 PUBLICATIONS **2,183** CITATIONS
SEE PROFILE

Ecological Informatics

Ecological Informatics

A Novel Automated Approach for Fish Biomass Estimation in Turbid Environments through Deep Learning, Object Detection, and Regression

Jansi Rani S V[a], Iacovos Ioannou[b,c], Swetha R[a], Dhivya Lakshmi RM[a] and Vasos Vassiliou[b,c]

[b]*Sri Sivasubramaniya Nadar College of Engineering, Chennai, Tamilnadu, 603110, India*

[b]*CYENS,, Nicosia 1016 Cyprus*

[c]*University of Cyprus, Nicosia 2109 Cyprus*

Elsevier use only: Received date here; revised date here; accepted date here.

Abstract

Estimating fish biomass is crucial in the fisheries sector, where traditional methods often harm fish through manual sampling and anesthetics. A non-invasive approach is introduced using underwater films to estimate fish biomass in turbid conditions. This study presents the "Aquatic WeightNet" dataset, targeting the Genetically Improved Formed Tilapia (GIFT) Tilapia species, and addresses the challenge of unclear images with preprocessing techniques like dehazing and Contrast Limited Adaptive Histogram Equalization (CLAHE). YOLOv8, a leading object detection model modified to accommodate the custom Aquatic WeightNet dataset's varied image sizes with five detection heads, P2 to P6, is employed, achieving a recall of 0.997 and a mean Average Precision (mAP) of 0.899 within the 50-95% Intersection over Union (IoU) range. Fish biomass estimation assesses depth, length, and width using regression models for calculation. A three-phase grid search identifies the most effective models, with the Extra Trees Regressor outperforming depth estimation with mean absolute error (MAE) of 0.63 and coefficient of determination (R^2) of 0.87 and the Random Forest Regressor for length and width (MAE of 0.01 and R^2 of 0.99). For biomass estimation, the Extra Trees Regressor again performs well (MAE of 0.004 and R^2 of 0.99), which is critical for determining optimal feed quantities to enhance aquaculture efficiency. This study emphasizes a non-invasive method to estimate fish biomass, optimizing the effectiveness and ecological sustainability of fish farming in murky waters through advanced detection algorithms and robust regression models.

Keywords: Fish Biomass; Object Detection; Regression; YOLOv8; mYOLOv8

1. Introduction

Fish biomass estimation is essential for the fisheries and aquaculture industries (Li et al., 2020). Manual methods of estimation can be strenuous for the fish. Thus, there is a need for non-intrusive methods of fish biomass estimation. Biomass estimation in a turbid environment also comes with several challenges. This is where deep learning can play a vital role. Training a deep learning model on large datasets of fish images can provide accurate biomass estimates without causing unnecessary strain to the fish and minimizes manual effort. Hence, the motivation of this research lies in providing an efficient, precise, and cost-effective method for estimating fish biomass in

turbid environments. The reliance on fishing and aquaculture is a prominent characteristic of numerous societies, playing a substantial role in bolstering the economic and social vitality of various nations and locations.

Fish plays a vital role in various domains such as trade, export, employment, food security, and nutrition due to its substantial contribution as a source of animal protein. The fisheries industry in India is considered a burgeoning sector, seeing rapid growth. In the fiscal year 2019-20, the Indian fishing sector produced 142 lakh tons, equivalent to around 15,652,820.6 tons. This notable output positioned India as the second-largest global fish producer, accounting for 7.56% of worldwide production. Furthermore, the fishing industry contributed significantly to the agricultural Gross Value Added, representing more than 7.28% of this economic indicator (Kumar, 2020). This industry demonstrates significant potential for expansion and the capacity to provide the basic requirements of marginalized groups. Fish biomass estimation is a widely employed and crucial technique in aquaculture. To effectively regulate stocking densities, optimize daily feeding practices, and determine the optimal harvesting timing, managers must consistently gather data on fish biomass. Nevertheless, the estimation of fish biomass is challenging without human intervention due to the fragile nature of fish and their capacity to navigate freely in an environment characterized by uncertain visibility, lighting conditions, and stability.

The utilization of fish biomass holds potential value in fisheries management as it provides a reliable approximation of the amount of feed that should be administered (Gaude et al., 2019). The estimation of fish biomass involves the multiplication of the average estimated weight of fish with the total number of fish present in each area, as described by Li et al. (2020). Insufficient feeding practices have been found to negatively impact the growth rate of fish, while excessive feeding can lead to wastage of feeding, increased waste output, and subsequent economic losses (Sun et al., 2016). Environmental conditions influence aquatic animals' feed intake (Sun et al., 2016).

Using fish biomass data can assist aquaculture enterprises in optimizing their facility investments and effectively managing water quality issues arising from excessive feeding practices. The quantitative evaluation of fish biomass is the fundamental basis for scientific fisheries management and conservation methodologies intended to ensure sustainable fish output. Hence, aquaculturists must possess the capability to quantify fish biomass accurately. This information can be used to support fisheries management and conservation efforts, assess the impacts of human activities on aquatic ecosystems, and evaluate the effectiveness of management strategies (Cochrane, 2002).

The evaluation of fish biomass has hitherto depended on intrusive, laborious, and time-consuming manual sampling techniques (Palomares et al., 2020). Therefore, developing a non-invasive, expeditious, and cost-effective approach is imperative. Multiple techniques have been suggested for the non-invasive assessment of fish biomass. Implementing measures to mitigate malnutrition risks is imperative, including underfeeding and overfeeding. Using fish biomass data assists aquaculture enterprises in optimizing capital investments in facilities and effectively managing water quality concerns arising from excessive feeding practices. The quantification of fish within the overall biomass is of utmost importance in understanding fish populations' trophic organization and reproductive capacity in coral reef ecosystems. Measuring fish biomass yields valuable insights into numerous factors, such as the overall health of fish populations, their growth rates, feeding behaviors, economic significance, and the quality of fish specimens. Fish biomass strongly correlates with numerous water factors, like turbidity, salinity, and mineral content (Matveev et al., 2013). The estimation of the biomass of the fish from underwater images in relation to turbidity has not yet been handled. The amount of feed to be dispersed can be calculated from the biomass of fish, which can help prevent overfeeding or underfeeding. Water is said to be turbid when cloudy or solid particles are suspended. Fish are

less efficient in turbid waters. A 5-6 ppm turbidity, or mg/L, is ideal for spawning (Ali et al., 2022). Accurately estimating fish biomass in aquatic environments is essential for many reasons, including monitoring the health of fish populations and managing commercial and recreational fisheries.

The most innovative aspect of this research is the introduction and utilization of the customized "Aquatic WeightNet" dataset, combined with advanced preprocessing techniques such as dehazing and Contrast Limited Adaptive Histogram Equalization (CLAHE), to enhance image clarity in turbid underwater conditions. This dataset, tailored specifically for the Genetically Improved Farmed Tilapia (GIFT) species, serves as the foundation for training and validating the modified YOLOv8 model (mYOLOv8). The addition of a fifth detection head to YOLOv8 significantly improves its ability to handle varied image sizes and detect fish accurately in murky waters, achieving a recall rate of 0.997 and a mean Average Precision (mAP) of 0.899 within the 50-95% Intersection over Union (IoU) range. Coupled with robust regression models like the Extra Trees Regressor and Random Forest Regressor, which are optimized through a three-phase grid search, the system offers highly precise biomass estimates (MAE of 0.004 and R^2 of 0.99). This non-invasive method not only minimizes harm to fish but also promotes ecological sustainability, optimizing feed management and enhancing the overall efficiency of aquaculture practices.

The novel approach introduced in this study offers a non-invasive method to estimate fish biomass using underwater films tailored explicitly for high turbidity conditions. Although previous works have used deep learning techniques to estimate fish biomass, they do so without including turbidity as a factor to estimate fish biomass. The novelty of this approach involves estimating fish biomass by including water turbidity as a factor affecting biomass. It also involves the modified YOLOv8 used for object detection, which shows improvement over the original model. This method contrasts with traditional, potentially harmful techniques. The research introduces the "Aquatic WeightNet" dataset tailored for the GIFT Tilapia species and utilizes advanced preprocessing techniques, such as dehazing and CLAHE, to enhance image clarity. The study modifies the YOLOv8 object detection model with architectural adjustments for dataset compatibility (hereinafter referred as mYOLOv8). The main purpose of this work is to develop a sophisticated method to accurately estimate fish biomass from underwater images of fish in turbid environments.

The objectives of this study include:

- Creating a custom specialized dataset of fish taken in a turbid environment.
- Finding the best model for object detection and making tweaks, if necessary, to the model to achieve better results.
- Finding suitable regression models for estimating fish biomass using depth estimation, finding the fish's length and width in centimeters, and finally estimating the biomass.

- The contributions made by this work encompass:

- **Introduction of a non-invasive fish biomass estimation method:** Traditional methods of fish biomass estimation often involve manual sampling and anesthetics, which can harm the fish. This contribution offers a way to estimate fish biomass without physically or chemically interfering with the fish, ensuring their health and safety.

- **Creation of the specialized "Aquatic WeightNet" dataset:** The dataset, tailored explicitly for the GIFT Tilapia species, provides a unique and robust collection of data necessary for developing and validating the proposed non-invasive methods. Such specialized datasets are essential for training and testing algorithms tailored for specific conditions and species.

- **Enhancement of image quality in turbid conditions through advanced processing:** Given the challenge of turbid underwater

conditions that often result in unclear images, advanced preprocessing techniques, such as dehazing and CLAHE, have been employed to improve the clarity and quality of these images.

- **Adaptation of YOLOv8 for robust fish detection:** YOLOv8, a leading object detection model, has been specifically adapted to cater to the unique needs of the dataset and the project goal. Architectural modifications ensure compatibility with the variable image sizes present in the dataset and enhance the detection of fish in diverse conditions.

- **Achievement of exceptional detection accuracy**: The adapted YOLOv8 (called m YOLOv8) model achieved outstanding performance metrics like a recall rate of 0.997 and a mean Average Precision (mAP) of 0.899 within the IoU range of 50-95%. This high level of accuracy is critical for reliable biomass estimation.

- **Development of regression models and optimization techniques:** Biomass estimation is not just about detection but also involves assessing parameters like the fish's depth, length, and width. The research employed regression models, such as the Extra Trees Regressor and Random Forest Regressor, optimized using a three-phase Grid Search methodology, ensuring precise and consistent biomass calculations. The grid search method uses hyperparameter tuning to find the most optimal hyperparameters of the model. Thus, by using this technique several techniques were evaluated and the best one of those was chosen at each step.

- **Promotion of ecological sustainability in aquaculture practices:** The study plays a pivotal role in reducing the need for invasive sampling techniques by devising a non-invasive method for estimating fish biomass. This approach promotes the well-being of the fish, leading to more ecologically sustainable and humane aquaculture practices. The accurate estimation of biomass also assists in optimizing feed quantities, further enhancing sustainability.

This study has a few limitations, which need to be considered. The dataset is specialized to capture images of fish underwater in a turbid environment. However, it does not focus much on extreme conditions such as the unavailability of light. Though the model can detect fish with some occlusion, it might not be very accurate in cases of extreme occlusion or overlap. In extremely saline conditions, for example, if the biomass of the fish is greatly affected by salinity, then the biomass may not be accurate since factors other than turbidity, such as salinity, are not taken into account, though in normal conditions, the other factors would have little effect on fish biomass.

The research hypothesis can be stated as follows: The non-invasive approach utilizing underwater videos, the YOLOv8 object detection model tailored with five detection heads (P2 to P6), and robust regression models for biomass estimation will yield more accurate and ecologically sustainable results in estimating fish biomass in turbid conditions compared to traditional invasive methods. This advanced approach is expected to contribute to the optimization of fish farming efficiency and ecological sustainability in turbid waters.

This research has significant implications for various stakeholders. The adoption of non-invasive fish biomass estimation techniques can revolutionize aquaculture practices globally, fostering sustainable and responsible fish farming. Accurate biomass estimation can optimize feed management, stocking decisions, and overall operational efficiency, benefiting aquaculture industries worldwide. By minimizing harm to fish through non-intrusive methods, this approach aligns with environmental sustainability targets and regulations.

Policymakers can consider integrating non-invasive biomass estimation techniques into aquaculture regulations to promote sustainable practices and minimize environmental impact. The advancement enables the development of low-cost, energy-efficient systems for fish farming, contributing to energy conservation and reduced emissions. Policymakers can align

regulatory frameworks with UN SDGs, specifically those related to responsible consumption and production, climate action, and life below water, by promoting the adoption of non-invasive biomass estimation practices.

Policymakers, industry stakeholders, and researchers can collaborate to integrate these advancements into regulations, standards, and practices, promoting sustainable and responsible aquaculture globally.

2. Literature Review and Background Work

2.1 Literature Review

This section provides the methods identified in the studies regarding fish biomass identification using image processing. Conventional methods of fish biomass estimation, such as netting and trawling, can be expensive and disruptive to the ecosystem. Underwater videos provide a non-invasive alternative to estimate fish biomass in various environments, including turbid waters. The problem of estimating fish biomass in turbid water using underwater videos is challenging due to visibility limitations. The water column can be cluttered with particles and sediment that obscure the view of the fish, making it difficult to detect and measure them accurately.

Some other biomass estimation techniques include machine-vision-based methods (Li et al., 2020), acoustics-based methods (Sun et al., 2016), environmental DNA-based techniques (Sun et al., 2016), and deep learning-based methods (Abinaya et al., 2022). Barbedo 2022 detailed employing computer vision and artificial intelligence-based fish monitoring and recognition methods. Biomass statistics are computed in sea cages (Liu et al., 2021). However, by applying image processing techniques, such as object detection and segmentation, it is possible to extract fish from the background and calculate their size and biomass. By using underwater videos, it is possible to detect the fish present and estimate their size, which can then be used to calculate the total biomass. Overall, formulating the problem of estimating fish biomass in turbid water using underwater videos provides a practical and non-invasive solution for monitoring fish populations and managing aquatic ecosystems. By carefully calibrating the camera and applying appropriate image processing techniques, it is possible to obtain accurate fish biomass estimates even in challenging environments.

The integration of ecological studies with semantic segmentation techniques can provide valuable insights into environmental monitoring and biodiversity assessment. For instance, satellite imagery combined with semantic segmentation algorithms can be used to monitor land cover changes, track deforestation, and assess habitat fragmentation. By classifying different land cover types (e.g., forests, water bodies, agricultural land) at a fine spatial resolution, researchers can analyze ecosystem dynamics and assess the impact of human activities on natural habitats.

The authors (Beveridge et al., 2008) explore the intricate connection between environmental sustainability and Asian aquaculture. The authors draw attention to aquaculture's dual function in meeting both supply and demand for environmental products and services. From the supply side, aquaculture helps to create environmental goods by sequestering carbon, creating habitats, and purifying water. These services are essential to sustaining biodiversity and the health of ecosystems.On the other hand, aquaculture operations also need healthy ecosystems, clean water, and appropriate habitats for their species. Aquatic ecosystems can suffer long-term damage and ecological deterioration due to unsustainable activities, including nutrient pollution, habitat destruction, and antibiotic misuse. The necessity of sustainable aquaculture methods that strike a balance between the supply and demand for environmental products and services is emphasized in the study. This entails implementing ecosystem-based strategies, embracing ethical farming methods, encouraging circular economies, and involving stakeholders in cooperative conservation initiatives. All things considered, (Beveridge et al.,2008).'s research underscores the vital role that Asian aquaculture plays in environmental sustainability and points the way toward more conscientious and sustainable aquaculture methods.

In field-based ecological studies, image analysis techniques such as semantic segmentation can aid in species identification and habitat mapping (Seyed et al,2022). By segmenting images captured by remote sensing or field surveys, scientists can demarcate different vegetation types, identify rare or invasive species, and monitor wildlife populations. This information is critical for conservation planning, ecosystem restoration, and biodiversity conservation efforts. Integrating semantic segmentation into ecological studies provides a powerful approach to understanding and managing ecosystems. Using advanced computational techniques, researchers can analyze large-scale ecological data sets more efficiently, leading to better-informed decision-making and sustainable environmental management strategies.

Several studies have been done on quantifying fish biomass in this new technological era. For enhancing the image, Ancuti et al., 2017 have specified a technique for image enhancement and restoration. It is based on merging two images produced by taking the original degraded image and applying colour correction and white balancing. The two images are fused using a multiscale fusion method. The various performance metrics that can be used are Mean Absolute Error (MAE), Mean Squared Error (MSE), Signal Noise Ratio (SNR), and Root Mean Squared Error (RMSE), etc. Some other techniques that are present include Wavelet Compensation and Image Dehazing, Contrast Limited Adaptive Histogram Equalization, Dark Channel Prior Method, Discrete Cosine Transform Method, etc.

An image-based mass estimation of free-swimming fish has been developed and tested (Lines et al., 2001). For estimating fish length, a stereovision-based method is designed and used (Shi et al., 2020). Junior et al., 2021 have used two different algorithms: a. J48, SVM (Support Vector Machines), and KNN and linear regression b—convolutional Neural Networks. J48 gave an accuracy of 58.2%, whereas CNN (Convolutional Neural Networks) gave an accuracy of 67.08% for estimating fish biomass. Biomass for fingerlings is estimated using the Deep Belief Network (Pache et al., 2022a), and for the live body biomass of Pintado Real fingerlings using supervised learning (Pache et al., 2022b). The non-intrusive methods (Li et al., 2020), such as stereovision, laser scanners, acoustics, eDNA, and resistivity counters, are summarized with advantages and disadvantages. The fish detection and tracking research for turbid underwater video (Gaude et al., 2019) uses background subtraction as a preprocessing technique to handle turbid images. Then, fish detection is done using a three-frame difference. Fish movement is tracked using an optimal estimation algorithm, the Kalman filter. Tolentino et al., 2020 have estimated the weight of Nile Tilapia using image processing and the Internet of Things (IoT).

V. F. Matveev et al., 2013 have discussed the relationship between biomass, turbidity and salinity. By analyzing the FB collected from the Logan River (main estuary) and the Albert River (tributary) located in Southeast Queensland, Australia, they have examined the relationship between biomass, turbidity, and salinity. The Albert River, which had a lower annual mean discharge, a higher salinity and pH, and a greater quantity of phosphorus, was shown to have a higher FB. Turbidity and salinity were two of the fifteen hydrological variables investigated, and they were found to be crucial factors in determining the fish biomass. The amount of biomass and water turbidity were shown to be negatively correlated.

Another system using Computer Vision was developed to estimate fish biomass after removing the fish's tail fin (Hao et al., 2022). Youssef Wageeh et al. (Wageeh et al., 2021) analyzed various algorithms - frame subtraction and optical flow and finally used the YOLO algorithm for object detection. YOLO (You Only Look Once) is an object detection algorithm in computer vision. It is a real-time object detection system that processes an entire image in one forward pass of a neural network. Dividing the input image into a grid of cells and predicting the presence of objects using bounding boxes in each cell is the objective of YOLO. This approach allows YOLO to process images quickly and make predictions in real-time. The combination of the MSR color correction algorithm and YOLO is designed to enhance the image. The research achieved an accuracy of 94% in detecting the fish. Deep learning and machine learning models are used for other areas in fisheries, such as fish species identification for conducting surveys (Yassir et al., 2023; Jalal et al., 2020), fish age estimation (Politikos et al., 2021; Bezer et al., 2022), finding length of the

fish and catch numbers (Palmer et al., 2022), categorizing squid species and freshness (Hu et al., 2020), recognizing sea cucumber (Xuan et al., 2023), tuna aggregation (Baidai et al., 2020), and tuna biomass estimation (Precioso et al., 2022).

Authors in (Tengtrairat et al., 2022) have developed a two-step biomass estimation algorithm for the biomass estimation of Tilapia in turbid water, where the first phase is Tilapia Detection, and the second phase is Tilapia Biomass Estimation. The fish were photographed underwater with a low-cost video camera. To identify and quantify the fish's visual dimensions, a Mask Region-based Convolutional Network model is trained. The experimental findings revealed an average weight inaccuracy of 30.30 (23.09) grams, a mean absolute error of 42.54 g, and an R^2 of 0.70. The images are taken from a turbid environment, but they have not designed any modifications concerning the turbid images.

The objective of Coro et al., 2021 is underwater fish detection, while the objective of this research is to estimate the biomass of fish. Object detection is the first module in this research. Following YOLO, the regression module is used for biomass estimation. Moreover, this research has a preprocessing module before YOLO object detection. The aim is to estimate fish biomass, not merely detect the fish. UDMOS, as described in Coro et al , 2021 is a combination of hardware and software components. It does real-time processing of underwater fish. This research, in contrast, is not real-time since recorded videos are processed and analyzed. Moreover, this research does not use a hardware system like the Raspberry Pi. UDMOS uses a combination of motion sensors, deep learning, and unsupervised learning techniques. Though the use cases vary between the two systems, one of the advantages of the UDMOS system is that it can capture a large amount of data in extreme conditions. It would definitely be useful to use all the data from UDMOS and use it in the proposed system after rigorous research.

One of the advantages of the proposed system over UDMOS is that a more advanced version of YOLO- v8 (and a modified version of v8) is used, while the UDMOS system uses v3 of YOLO. Moreover, the sheer amount of processing done in the proposed system would be more suitable for fish biomass estimation.

The precision and recall values are common to both systems. However, since the proposed system is for a single species (GIFT Tilapia) and the UDMOS system is for any aquatic animal in a very complex environment and uses complex equipment, it would not be practical to compare the results. 75.8% and 82.5% are the precision and recall achieved in the UDMOS system, while the proposed system achieves much higher values. Since there are too many different variables, a fair comparison is not possible.

This section briefly reviews the literature survey studied for this research. Table 1 summarizes the main methods and their performance metrics. Depending on specific needs, each approach can be expanded, and the methodologies and results can be delved deeper into.

Table 1. Comparison of existing models and proposed models

Method / Study	Approach / Technique	Performance Metrics (if applicable)

Abinaya et al., 2022	Deep learning-based methods for biomass estimation.	Yolov4: 91.52% accuracy
Liu et al., 2021	Biomass statistics in sea cages using image processing techniques.	Average Mean Geometric Boundary: 0.52 IoU
Junior et al., 2021	a. Algorithms: J48, Support Vector Machine, K Nearest Neighbours, and linear regression for biomass estimation. b. Convolutional Neural Networks for biomass estimation.	J48: 58.2% accuracy, CNN: 67.08% accuracy
Pache et al., 2022a	Estimation of biomass for fingerlings using Deep Belief Network.	Deep Belief Network: 0.7 R^2
Pache et al., 2022b	Biomass estimation of Pintado Real fingerlings using supervised learning.	Linear Regression: 0.76 R^2
Tolentino et al., 2020	Weight estimation of Nile Tilapia using image processing and IoT.	Regression: 2.82 Mean Percentage Error
Hao et al., 2022	A computer vision system will be used to estimate fish biomass after removing the fish's tail fin.	Partial Least Square: 7.10g RMSE, 5.36g MAE, 8.46% MaxRE
Wageeh et al., 2021	Used object detection algorithms, including frame subtraction, optical flow, and YOLO. Enhanced the image using the MSR color correction algorithm.	94% accuracy in fish detection using YOLO.
Tengtrairat et al., 2022	Developed a two-step biomass estimation algorithm for Tilapia. Used Mask RCNN model to identify and quantify fish's visual dimensions.	Average weight inaccuracy of 30.30 grams, MAE: 42.54 g, R^2: 0.70

Our Study	Utilized underwater films for fish biomass estimation. Preprocessing techniques like dehazing effects and CLAHE were used. Employed YOLOv8 with architectural alterations. Regression models like Extra Trees Regressor and Random Forest Regressor were used.	mYOLOv8 99.98% Recall rate: 0.997, mAP: 0.899 (IoU 50-95%). Depth (Extra Trees Regressor): MAE: 0.63, R^2: 0.87. Length & Width (Random Forest Regressor): MAE: 0.01, R^2: 0.99. Biomass (Extra Trees Regressor): MAE: 0.004, R^2: 0.99.

2.2. *Review and Reasoning for the approaches used*

Object detection algorithms play a vital role in detecting fish, while biomass estimation methods provide valuable insights into population dynamics. Accurate estimation of fish biomass is essential for fisheries management and ecological studies, often relying on image analysis techniques for non-invasive measurements. However, underwater imaging poses challenges such as uneven illumination and low contrast, necessitating effective image preprocessing techniques.

Traditional image enhancement methods, such as histogram equalization, contrast stretching, and filtering, are commonly used to improve image quality. While these methods can enhance certain aspects of images, they may not adequately address the specific challenges encountered in underwater environments, including variable lighting conditions and low contrast in shadowed areas. By dividing the image into smaller regions and independently equalizing the histogram of each region, CLAHE preserves both global and local contrast, making it particularly effective for enhancing images with uneven illumination and low contrast. This adaptability to local image characteristics is crucial for accurately identifying and measuring fish in underwater images.

Several state-of-the-art object detection algorithms were reviewed, including Faster R-CNN, SSD, and YOLOv8. While all algorithms have
effectively detected objects, YOLOv8 stands out due to its superior speed and accuracy. YOLOv8's single-stage architecture allows for real-time processing without sacrificing detection performance, making it well-suited for applications requiring rapid and accurate object detection, such as fisheries monitoring.

Various methods exist for estimating fish biomass, including length-weight relationships, acoustic techniques, and regression analysis. Regression analysis emerges as the most reliable and versatile method for biomass estimation. Regression models can account for various factors influencing fish biomass, such as length, weight, and environmental variables. Additionally, regression analysis allows for developing predictive models that can accurately estimate biomass across different species and habitats.

2.3. *Background Work*

This section provides background information regarding the utilized approaches.

1. **Dehazing Effects:** Image dehazing is the process of removing haze from images to increase their visibility. Haze is a phenomenon caused by atmospheric conditions where light gets scattered by the presence of tiny particles, resulting in a decrease in the contrast of the captured image. Mathematically, the hazy image can be represented by a linear model:

$$I(x) = J(x)t(x) + A(1 - t(x)) \qquad (1)$$

where,

I(x): observed intensity
J(x): scene radiance (what is needed)
t(x): medium transmission
A: global atmospheric light

Algorithm: Various algorithms are available for dehazing, with the Dark Channel Prior being

the most prevalent. This method operates on the notion that within the majority of non-sky regions, there are pixels in at least one color channel that exhibit extreme darkness or exceptionally low intensities.

2. **CLAHE (Contrast Limited Adaptive Histogram Equalization)**: CLAHE is an advanced method of histogram equalization that enhances the contrast of images. Instead of considering the full image, CLAHE divides the image into smaller blocks and applies histogram equalization to each block. This method can enhance the local contrast of an image and limit the amplified noise. The 'contrast limiting' step ensures the noise is not enhanced.

$$O(x,y) = Equalize(I(x,y)) \quad (2)$$

where,

O(x,y): output image
I(x,y): input image

Algorithm: For every small block:
1) Compute the histogram.
2) Apply contrast limiting.
3) Compute the equalization transform.
4) Apply the transform to the block.

3. **YOLOv8 with Architectural Alterations**: YOLO (You Only Look Once) is a real-time object detection system. YOLOv8 is a more advanced version, which means it has improved accuracy and speed. Architectural alterations refer to changes made to the neural network architecture to better suit a specific task or dataset.

$$P(C,B|I) = NN(I) \quad (3)$$

$P(C,B|I)$: $Probability\ distribution$ over classes C and bounding boxes B given an input image NN: Neural network function.

Algorithm: The fundamental idea behind YOLO is to divide the image into a grid. Predictions for bounding boxes and class probabilities are made for each grid cell. The architectural alterations might involve adjusting the number of layers and filters or introducing new components to the neural network.

4. **Extra Trees Regressor:** The Extra Trees Regressor is an ensemble learning method. It operates by constructing many decision trees at training time and outputs the average prediction of the individual trees to regress. It is "extra" because it randomly chooses each feature's threshold instead of searching for the best possible thresholds like a random forest. Equation 4 talks about the prediction of extra trees regressor. If 'N' is the number of trees, and Ti is the prediction of the ith tree, then the prediction of the extra trees regressor is given by the equation.

$$f(x) = \frac{1}{N}\sum_{i=1}^{N} \quad T_i(x) \quad (4)$$

where,

$f(x)$: $Prediction\ for\ input\ x$
N: $Number\ of\ trees$
T_i: $Prediction\ of\ the\ ith\ tree$

Algorithm:
1. Draw a bootstrap sample of the dataset.
2. Randomly select features for each tree node and choose the best split based on these features.
3. Repeat the process for each new node until a stopping criterion is met.
4. Aggregate predictions from all the trees for the final prediction.

5. **Random Forest Regressor**: Random Forest Regressor is another ensemble learning method used for regression (and classification). It builds multiple decision trees and merges them together to get a more accurate and stable prediction. A key distinction from the Extra Trees Regressor lies in

its approach to selecting splitting thresholds, although the formulation remains identical to that of the Extra Trees Regressor.

Algorithm:

1. Draw a bootstrap sample of the dataset.
2. Grow a decision tree from the bootstrap sample. At each node: a. Randomly select a subset of features. b. Determine the best split based on these features.
3. Repeat the process for each new node until a stopping criterion is met.
4. Aggregate predictions from all the trees for the final prediction.

The following section describes the materials and methods involved in the system, including dataset details, proposed methodology, training procedure, and performance evaluation parameters—the section following that tabulates the results and discussion, followed by the conclusion and references.

3. Methods

3.1 Dataset

The Aquatic WeightNet dataset consists of fish images and videos that were custom-prepared. The fish images were collected from Dr. M.G.R. Fisheries College and Research Institute, Ponneri, Chennai, Tamil Nadu, India. The images were taken with a phone camera. The specifications for the camera used are as follows. The advanced dual camera system has 12MP Wide and Ultra Wide cameras. The Wide lens has a 26mm focal length (35mm equivalent). Its field of view is approximately 79°. A baseline distance of a few millimeters to a centimeter between the lenses enhances depth-sensing capabilities. The camera was mounted in a fixed position at a distance of 0.5m from the fish tank, and the photographs were captured. 30 fish of GIFT Tilapia species were taken in a tank, and their length, width, and biomass were measured. Videos of fish were taken and converted to frames. This created a total of 5325 images, which were split into 70% for training, 10% for testing, and 20% for validation. All these images were annotated using CVAT.ai and uploaded to be managed on Roboflow.

3.2 Proposed Methodology

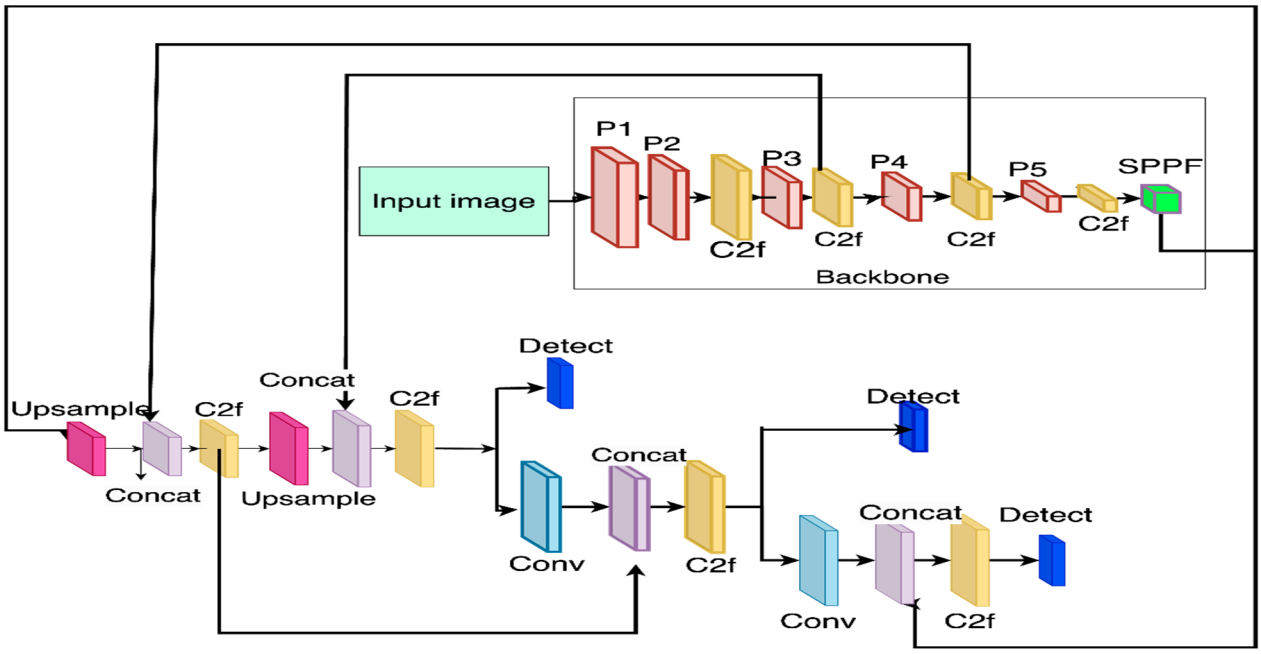
The proposed methodology aims to develop a biomass estimator for fish in a turbid environment using YOLOv8 and regression modules. The system flow is defined as shown in Figure 1. Underwater videos of fish that have been collected are processed to get individual frames or fish images. The images are pre-processed using image enhancement and restoration techniques. These images are then passed to the object detection algorithm, YOLOv8 and mYOLOv8 (new, proposed architecture of YOLOv8).

The object detection algorithm outputs bounding boxes around each fish in every frame. The bounding box around any fish is used in pixels to calculate the length and width of the fish. The biomass has to be estimated from the length and width in pixels. This is done with the help of regression models. The conversion of length and width in pixels (as calculated by the bounding box) to the dimensions of the fish (length and width in centimeters) requires some information about the distance between the camera and the fish. This is called the depth information. The length (or width) in pixels does not give the actual size of the fish. This is because the same fish, as seen from close to the camera, will appear larger, resulting in larger pixel measurements. From afar, the fish might appear smaller. This is why the depth information is important. Thus, the output of the object detection model is used as input for the regression models.

From the object detection step output, the length and width of the fish in pixels are derived. There are three regression models. The first of these is the depth estimator. The depth estimator model takes as input the length and width in pixels and gives the depth as output. That is, the length and width of the fish in pixels along with the depth, which is the ground truth data, is used to train this depth estimator model. Thus, the model learns how to predict depth given the input length and width of the fish in pixels. This is how the depth estimator model outputs depth or the distance between the camera and the fish, when it is provided with the input of length and width in pixels from the object detection step. The second model is the pixel-to-centimeter converter, which takes as input

the length and width in pixels, along with depth from the output of the previous step, and outputs length and width in centimeters. Fish biomass depends on factors such as fish length, width, and turbidity of water. The final model tries to establish a relationship between these factors. The last model, the tilapia weight estimator, takes the length and width in centimeters and turbidity as the input and gives the estimated biomass as output. There are many regression models available. To find the one that best suits each step, grid search (Rao & Vinod, 2019) is employed. It uses hyperparameter tuning to find the best regression technique for each. Finally, the biomass is fed to a feed estimator, which outputs the amount of feed required since the amount of feed depends on the biomass of the fish.

Fig. 1. Proposed Architecture



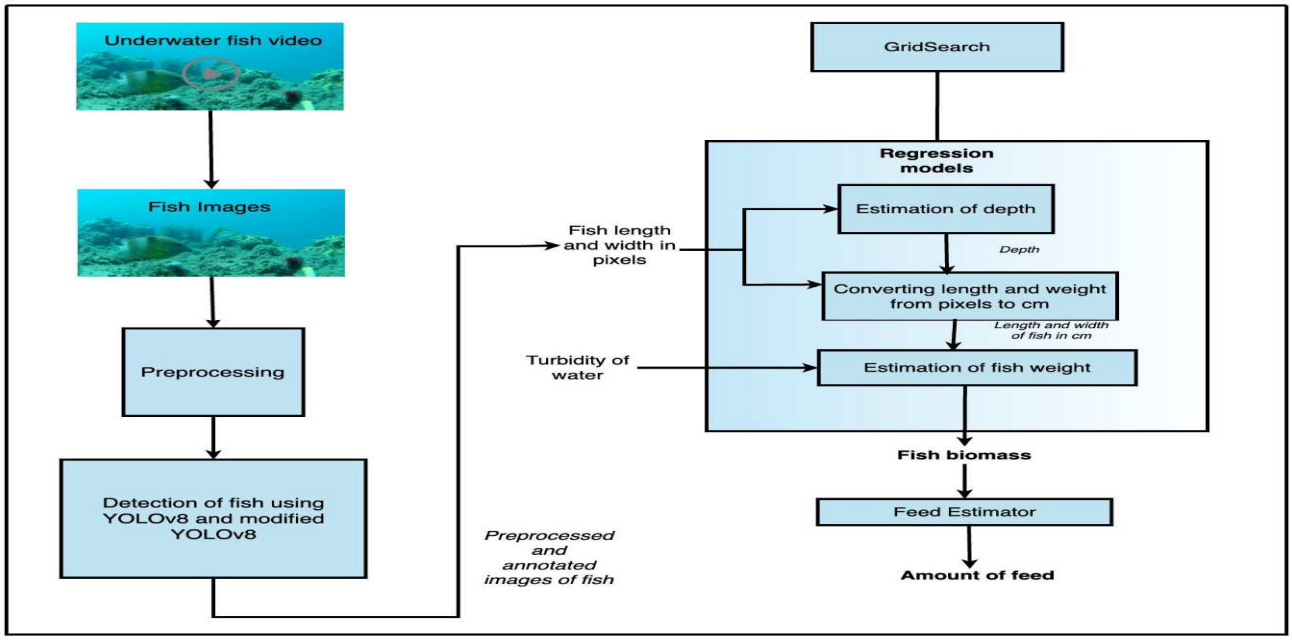


Fig. 2. YOLOv8 original architecture

3.2.1 Image Preprocessing

Images were pre-processed to remove the murkiness and haze in the images. Four Techniques of image enhancement were used:

- **White balance followed by Gamma Correction and Sharpening:** Image Fusion using Wavelet Transform is done to produce an enhanced image.
- **Dehazing:** It reduces atmospheric haze in color or grayscale image.
- **Denoising:** It does dehazing followed by edge-preserving smoothing on an image.
- **Contrast Limited Adaptive Histogram Equalization (CLAHE):** This adaptive histogram equalization limits contrast amplification to reduce noise amplification.

3.2.2 YOLOv8 Architecture

• Ultralytics undertook the development of YOLOv8. The development process involved implementing enhancements to the existing YOLOv5 model. The field of architecture can be categorized into two components: the backbone and the head. The backbone utilized in this context is a modified iteration of CSPDarknet53. The cranium's structure consists of multiple intricate layers, followed by fully connected layers. The backbone consists of five convolutional layers labeled P1 to P5, as depicted in Figure 2. The Sigmoid Linear Units(SiLU), or swish activation function, is employed in all convolution layers. The Sigmoid Linear Unit is an activation function commonly used in neural networks, promoting smooth gradients and efficient learning by combining the properties of sigmoid and linear activations. Layers P2 through P5 are subsequently connected to a C2f (convolution to fully connected) layer, followed by a CSP Bottleneck consisting of two convolutional operations. The present model exhibits output in the P3-P5 layers, accommodating strides ranging from 8 to 32. The subsequent layer in the network architecture is the Spatial Pyramid Pooling Fast (SPPF) layer, which represents an enhanced iteration of the original Spatial Pyramid Pooling(SPP). Spatial pyramid pooling is a technique in computer vision that partitions an image into sub-regions and pools features from each sub-region at multiple scales, enabling efficient and effective representation learning for varying spatial resolutions. Spatial

pyramid pooling is employed to construct a representation of fixed length. The proposed method performs a single computation of feature maps for the entire image and subsequently applies pooling operations to build a fixed-length representation. The pooling operation is executed by employing the Max Pool function. The output of the SPPF layer is subjected to up-sampling and then combined with the C2f output of the P4 backbone by concatenation. An additional layer of C2f follows it. The resulting output is subject to a series of comparable procedures, including up sampling, concatenation with the C2f of the P3 backbone, and, finally, a C2f operation. This is one of the components of the Detect system. The C2f layer undergoes a series of further convolution layers. Subsequently, it is concatenated with the P4 head and sent through another C2f layer. This serves as the second iteration of the Detect head. Subsequently, the preceding output undergoes a comparable series of convolutions, followed by concatenation with the output of the Spatial Pyramid Pooling Fusion (SPPF) and another Convolutional Layer 2 (C2f), resulting in the formation of the third detect head.

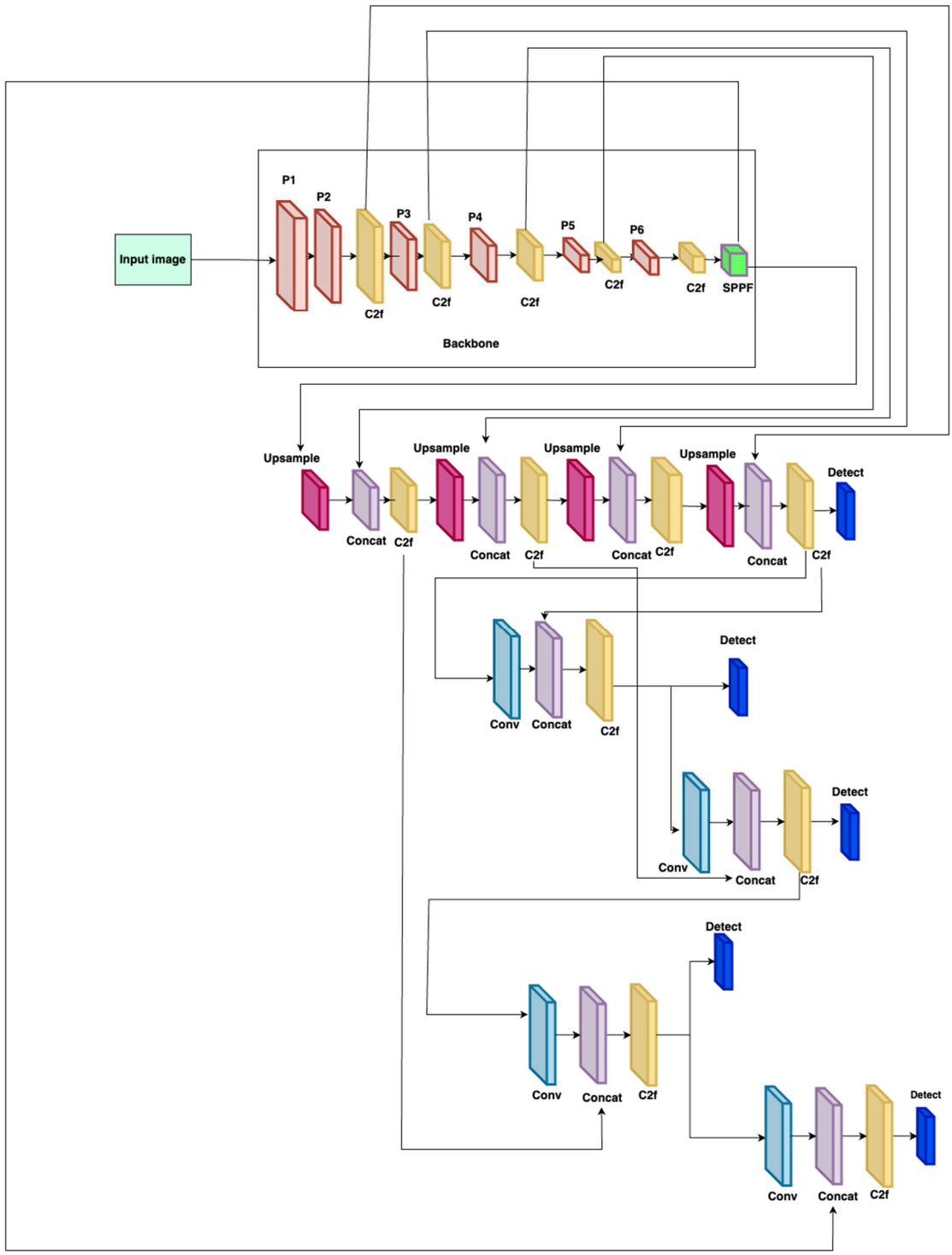


Fig 3. mYOLOv8 Architecture - five detect heads

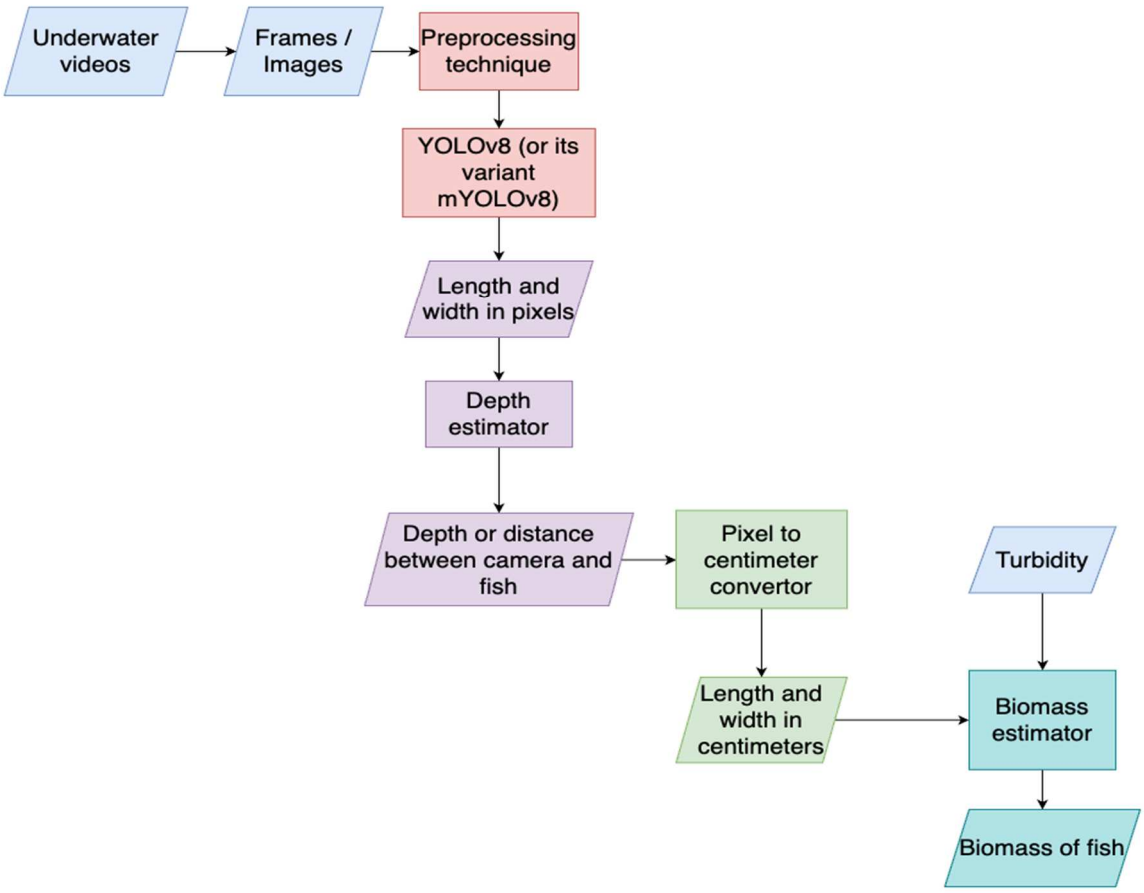


Fig. 4. Flowchart of the process

3.2.3 Proposed mYOLOv8 architecture with five detection heads

The proposed modification to the YOLOv8 consists of five detection heads, P2 to P6. Ultralytics has already given a variant of YOLOv8 with four detection heads, which extends the original architecture with a P6 convolution layer and makes it an additional detection head. This architecture is then extended by using the P2 layer for detection- this is the proposed architecture. Minor changes are made to include P2 for detection, as shown in Figure 3. A new Upsample-Concat-C2f block and a Conv-Concat-C2f block are added. P2 detect head is added. This can be helpful to support large and small image sizes. Note that in Figure 4, a flowchart is provided to give an overview of the entire process. It starts with taking the video input, converting it into images, preprocessing, followed by object detection, and then the three regression models, namely the depth estimator, the pixel to centimeter convertor, and the biomass estimator, at the end of which is the result, the biomass of the fish.

3.2.4 Biomass Estimation

Estimating the fish's biomass requires the fish's real-life dimensions- width and length in cm as input. Videos of the fish are captured, and the dimensions of the objects are estimated from them. Estimating the real-life dimensions of an object using a 2D image requires the following parameters as input:

1. Length and width in pixels.
2. Distance from the camera to the object.

3. Focal length of the camera used.
4. An object placed close to the camera is more prominent in the image than the same object placed far away; therefore, the focal length of the camera and the distance between the camera and the object are required.

It is rare to know the distance between the camera, the object, and focal length while using the biomass estimator. Thus, there is a compelling need to estimate these unknown parameters from the image provided. Regression is the best solution when there is a need to establish a relationship between various features. This led to the birth of the models:

1. Depth Estimator estimates the distance between the fish and the camera given the length and width in pixels.
2. Pixel to Centimeter Converter establishes a relationship between length and width in pixels and length and width in cm with the help of depth estimated by the previous model as a feature.
3. Biomass Estimator, which estimates the fish biomass in grams with the help of length and width in cm estimated by the previous models.

Algorithm 1. Overview of the proposed system

1. Convert video input into images.

2. Preprocessing techniques: To enhance turbid images.
   2.1. Preprocessing technique 1: White balance followed by Gamma Correction and Sharpening
      2.1.1. Compensate red channel attenuation.

$$I_{rc}(x) = I_r(x) + \alpha.(I_g^- - I_r^-).(1 - I_r(x)).I_g(x)$$

where,

I_r represent the red and green color channels of the image *I*

I_r^-, I_g^- denote the mean value of I_r

α denotes a constant parameter (α = 1)

      2.1.2. Compensate blue channel attenuation wherever blue is strongly attenuated.

$$I_{bc}(x) = I_b(x) + \alpha.(I_g^- - I_b^-).(1 - I_b(x)).I_g(x)$$

where,

I_b represent the blue and green color channels of the image I

α = 1

      2.1.3. Perform multi-scale fusion after gamma correction and sharpening.

   2.2. Preprocessing technique 2: Dehazing
   2.3. Preprocessing technique 3: Denoising
   2.4. Preprocessing technique 4: CLAHE
      2.4.1. Divide the image into segments or tiles.
      2.4.2. Perform histogram equalization on each tile using a predefined clip limit.
      2.4.3. Bilinear interpolation is done to stitch the tiles together.
   2.5. Choose the technique that gives the best results and proceed.

3. Detect the length and width of the fish. This is done by using the object detection algorithm YOLOv8. Two variations of the YOLOv8 algorithm are used. The original YOLOv8 and the modified YOLOv8.

    3.1. Original YOLOv8.
        3.1.1. The image is passed through 5 convolution layers P1 to P5. A C2f layer follows all layers except P1. The C2f layer is a CSP Bottleneck with two convolutions.
        3.1.2. SPPF or Spatial Pyramid Pooling (Fast) is performed, which pools the features and produces a fixed-length output.
        3.1.3. The image is passed through a series of Up-sample, Concatenation, and C2f layers.
        3.1.4. It is then passed through a detection head corresponding to levels P3-P5. This is the layer in which detections are made.

    3.2. Modified YOLOv8 (mYOLOv8).
        3.2.1. The image is passed through 6 convolution layers P1 to P6. A C2f layer follows all layers except P1. The C2f layer is a CSP Bottleneck with two convolutions.
        3.2.2. SPPF or Spatial Pyramid Pooling (Fast) is performed, which pools the features and produces a fixed length output.
        3.2.3. The image is passed through a series of Up-sample, Concatenation, and C2f layers. These layers are changed according to the new convolution layer, added in step 3.2.1.
        3.2.4. It is then passed through a detection head corresponding to levels P2-P6. The increased number of detection heads helps detect smaller and larger images.

    x_{w_pix}, x_{h_pix} are the width and height of the detected fish from step 3

4. Estimate depth using the length and width of fish in pixels.

$$y_{depth} = f(x_{w\_pix}, x_{h\_pix}, a_{w\_pix}, a_{h\_pix}) + e_{depth}$$

where x_{w_pix} and x_{h_pix} represent the width and height of the detected fish in pixels and a_{w_pix} and a_{h_pix} are constants.

5. Convert the length and width of fish from pixels to centimeters.

$$y_{l\_cm}, y_{w\_cm} = f(x_{w\_pix}, x_{h\_pix}, y_{depth}, a_{w\_pix}, a_{h\_pix}, a_{depth}) + e_{cm}$$

where y_{l_cm} $represents\ the\ length\ in\ cm\ and$ y_{w_cm} represents the width in cm.

x_{w_pix} and x_{h_pix} represent the width and height of the detected fish in pixels

y_{depth} represents the depth.

a_{w_pix} , a_{h_pix} , a_{depth} are constants.

6. Estimate fish biomass using the previous model's output and include turbidity as a factor.

$$y_w = f(x_{w\_pix}, x_{h\_pix}, y_{depth}, y_{l\_cm}, y_{w\_cm}\ a_{w\_pix}, a_{h\_pix}, a_{depth}, a_{l\_cm}, a_{w\_cm}, turbidity) + e_{cm}$$

where y_{l_cm} $represents\ the\ length\ in\ cm\ and$ y_{w_cm} represents the width in cm.

x_{w_pix} and x_{h_pix} represent the width and height of the detected fish in pixels.

y_{depth} represents the depth.

a_{w_pix} , a_{h_pix} , a_{depth} are constants.

$turbidity$ represents the turbidity of the environment.

3.3 Training Procedure

The Aquatic WeightNet dataset consists of fish images and videos of 30 fish of GIFT Tilapia species. It has a total of 5325 images. It is split into 70% training, 20% validation, and 10% testing. YOLO automatically resizes the images, so there is no need to adjust the image sizes. The dataset was captured manually with a mobile camera. It is not publicly available, but the authors can share it on demand. In the first step of object detection, the four different models were trained using the Aquatic WeightNet dataset. Each model was trained for 50 epochs. The bounding boxes are obtained from the object detection module, which calculates the length and width. This is passed to the depth estimator model, trained using ground truth. The output of this model is the depth, or the distance between the camera and the fish. The pixel-to-centimeter conversion model takes input length and width in pixels and depth (from the depth estimation model) and outputs the length and width in cm. The last model is the biomass estimator. It takes input length and width in cm and turbidity and outputs biomass in mg.

bounding boxes over their union is called intersection over union, or IoU. The IoU is set at a specific limit, either a number or a range. IoU set at 50% aids in the computation of mAP50. For computing mAP50-95, IoU fixed at a range of 50% to 95% with 0.05 increments is used. In this situation,

3.4 Performance evaluation parameters

Precision, Recall, and mAP (mean average precision) are the performance metrics used to compare object detection performance.

Precision: Precision provides information about the projected segment's accuracy.

Recall: Recall measures the ability of a model to identify all relevant instances of a particular class.

mAP- Mean average precision: The measure of choice for object detection algorithms is the mAP score. The region beneath the precision-recall curve is AP (Average Precision). A series of precision-recall curves is plotted with increasing difficulty levels in IoU and compute mAP. The overlap ratio between the actual and predicted
the mAP for each value in the range is computed. The AP for each class is determined as the area under the curve. mAP is calculated as the mean of APs.

The coefficient of determination or R^2 score, MAE - Mean Absolute Error, and RMSE-Root

Mean Square Error are the metrics used for the regression module.

Coefficient of determination or R^2 score: It is used to assess the performance of a linear regression model. The input-independent variable(s) predicts the variation of the output-dependent attribute. Depending on the ratio of the total deviation of outcomes given by the model, it is used to assess how effectively the model reproduces observed results.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5)$$

where,
SS_{res}: Sum of squares of the residual errors.
SS_{tot}: Total sum of the errors.

Residual errors represent the differences between the observed values of the dependent variable and the values predicted by the regression model. These residuals quantify the deviation or the "error" between the actual data points and the values estimated by the regression model.

MAE - Mean Absolute Error: MAE calculates the average difference between the calculated and actual values. It serves as a machine learning evaluation metric for regression models. It calculates the errors between the model's predicted and actual values. It is used to forecast the machine learning model's accuracy.

$$MAE = \frac{1}{M} * \sum |O_i - P_i| \quad (6)$$

where,
O_i: Actual value for the i^{th} sample.
P_i: Calculated value for the i^{th} sample.
M: Total number of samples.

RMSE-Root Mean Square Error: RMSE measures how well a regression line fits the data points. RMSE can also be construed as Standard Deviation in the residuals.

$$RMSE = \sqrt{\frac{\sum(Predicted - Actual)^2}{M}} \quad (7)$$

where,
Predicted: The predicted value for the i^{th} sample
Actual: The observed(actual) value for the i^{th} sample

M: Total number of samples.

4.1 Details of validation from the Aquatic WeightNet Dataset

915 GIFT images from Aquatic WeightNet are considered to validate the developed model, and they are detailed in Table 2.

Table 2
Parameters Values for validation set.

Fish Species	Biomass range	Image resolution	No. of images
Genetically Improved Farmed Tilapia (GIFT) (Scientific Name: Oreochromis niloticus L.)	10-126g	1920p*1080p	915

4. Results and Discussion

This study aimed to verify the validity of a previously constructed model utilizing photographs obtained from a specific dataset. The model's performance was evaluated based on data that had not been encountered before. This practice guarantees that the model's performance is not limited to the specific dataset it was trained on but rather extends to novel and unobserved scenarios. The validation photos are obtained from a proprietary dataset called "Aquatic WeightNet." To ensure the validity of the results, a significant number of 915 photos were utilized, indicating a comprehensive validation procedure encompassing a diverse array of cases. The photographs presented depict the Genetically Improved Farmed Tilapia (GIFT), which is a well-known abbreviation for this particular fish species. Scientifically, these fish are classified as "Oreochromis niloticus L." The biomass, denoting the aggregate mass of the fish depicted in the photos, exhibits variability,

encompassing fish specimens weighing between 10 and 126 grams. This implies that the model undergoes evaluation concerning a wide variety of fish sizes, thereby examining its capacity to adapt and its level of precision over this continuum. In addition, the photographs include a resolution of 1920 pixels by 1080 pixels, which guarantees a high level of clarity and intricate detailing. The provided data offers a comprehensive overview of the validation dataset, demonstrating the diverse scenarios utilized to evaluate the model's performance. This enhances the credibility and practical relevance of the model's findings in real-world situations. Table 2 specifies the parameter values for validating the GIFT fish image base, namely the species name, biomass range, image resolution, and number of images.

4.2 Experimental Results for the proposed method using Aquatic WeightNet dataset

The results of the experiments are presented in this section. As mentioned earlier, the Aquatic WeightNet dataset is considered. The results before and after preprocessing are presented with performance metrics such as precision, recall, and mAP. The outputs, along with the losses, are also presented. The results of the proposed novel methodology, mYOLOv8, are compared with those of YOLOv8. For estimating fish biomass, a suitable regression model is found using the grid search technique.

Hyperparameters are variables used to determine the model's ideal parameters. Grid search inputs all hyperparameter combinations one by one into the model to find the best values for a given model. After object detection, the bounding box coordinates are used to find the length and width of each fish in pixels. Three regression models are defined to estimate length from the pixel values of dimensions. At each step, the grid search finds the best regression method by examining all possible values individually. The last regression model is the biomass estimator, which outputs the fish biomass.

Finally, the obtained fish biomass is compared with the existing method.

4.2.1 Experimental Results for the proposed method before and after preprocessing

Table 3 showcases the outcomes of the YOLOv8 object detection model across different preprocessing circumstances. Commencing with the evaluation of the performance in its raw state, it is apparent that YOLOv8 demonstrates a robust baseline, as evidenced by Precision, Recall, and mAP50 values, all surpassing 0.96. Nevertheless, the mean Average Precision (mAP) for IoU thresholds ranging from 50% to 95%, known as mAP50-95, exhibits a value of 0.81, indicating the model's overall efficacy across a wider spectrum. A slight enhancement is found when employing the integrated preprocessing technique of white balance, gamma correction, and sharpening, commonly employed to augment images' brightness, contrast, and sharpness. The precision metric increases to 0.97, while the mean average precision at 50% overlap (mAP50) improves to 0.979. However, the recall metric remains constant at 0.96. Interestingly, the mean average precision (mAP) score from 50% to 95% intersection over union (IoU) remains stable, indicating that these strategies provide minimal advantages, particularly when considering wider IoU thresholds.

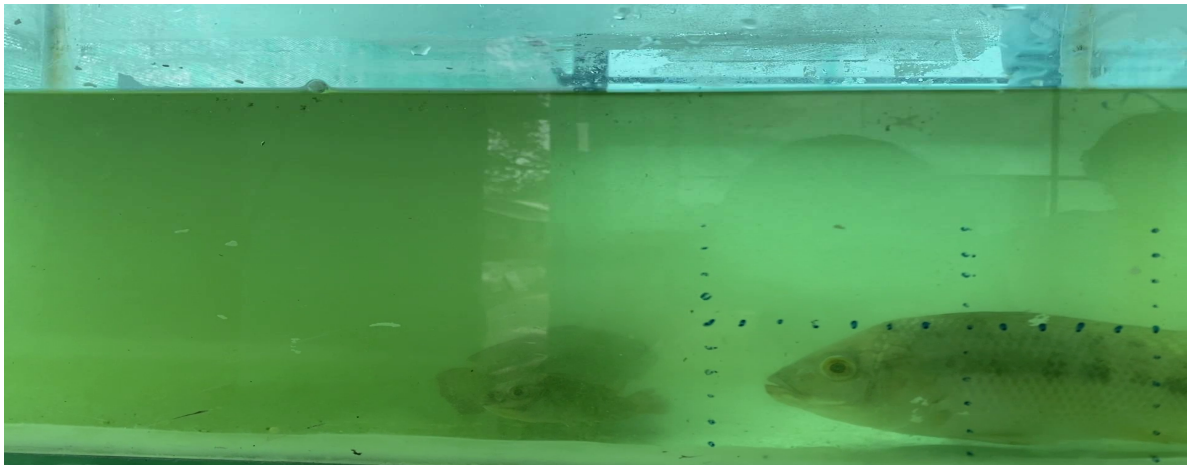
The process of dehazing, particularly in the domain of underwater imaging, yields substantial improvements. Following the dehazing preprocessing technique, the values of Precision and Recall show a notable increase, reaching a level of 0.98. Additionally, the mean average precision at 50% overlap (mAP50) exhibits further enhancement, achieving a value of 0.99. It is worth mentioning that the mean Average Precision (mAP) for IoU thresholds ranging from 50% to 95% exhibits a notable rise to 0.83. This indicates that the dehazing technique significantly enhances the model's performance across various IoU thresholds. The process of denoising, which aims to eliminate undesired random variations in brightness or color, gives a distinct perspective. Using denoising techniques, the obtained precision value is 0.97, while the recall value is 0.96. These values closely resemble the baseline statistics. The mean average precision at 50 (mAP50) demonstrates a little increase, reaching a value of 0.98. However, the

mean average precision from 50 to 95 (mAP50-95) remains steady at 0.81. This suggests that although denoising can slightly improve detection in certain situations, its overall effect on performance is still limited. Finally, the employment of Contrast Limited Adaptive Histogram Equalization (CLAHE), a technique specifically developed to equalize image contrast, has a significant impact. When using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique, the precision metric achieves a value of 0.97, while the recall metric significantly improves to 0.98. Both the mAP50 and mAP50-95 measures have been improved, achieving scores of 0.99 and 0.84, respectively. The observed substantial improvement, particularly in the mAP50-95 metric, suggests that the utilization of CLAHE significantly enhances the performance of the YOLOv8 model across a wide range of Intersections over Union (IoU) thresholds when compared to alternative preprocessing techniques.

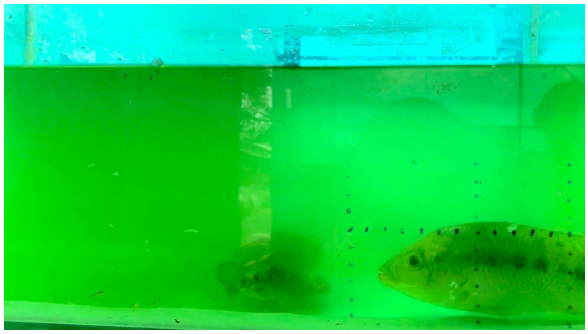
Table 3
Results of YOLOv8 after the four preprocessing methods

Technique	**Precision**	**Recall**	**mAP50**	**mAP50-95**
Without preprocessing	0.96	0.96	0.97	0.81
White balance, gamma correction and sharpening	0.97	0.96	0.979	0.81
Dehazing	0.98	0.98	0.99	0.83
Denoising	0.97	0.96	0.98	0.81
CLAHE	0.97	0.98	0.99	0.84

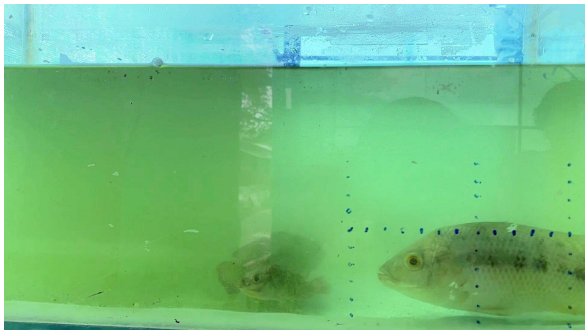
Figure 5 shows the images before and after preprocessing using the four preprocessing techniques. It highlights the importance of preprocessing approaches in facilitating object detection. However, it is important to note that the effectiveness of these strategies may differ across different scenarios or applications. In the specific domain of underwater imaging examined in this work, the preprocessing techniques of Dehazing and CLAHE have been identified as the most efficacious methods, resulting in significant enhancements across the many metrics that were studied.



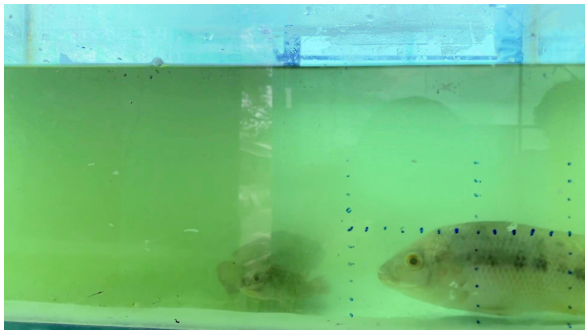
a) Original image



b) Technique 1: White balance followed by gamma correction and sharpening



c) Technique 2: Dehazing



d) Technique 3: Denoising

e) Technique 4: CLAHE

Fig. 5. Output of preprocessing techniques

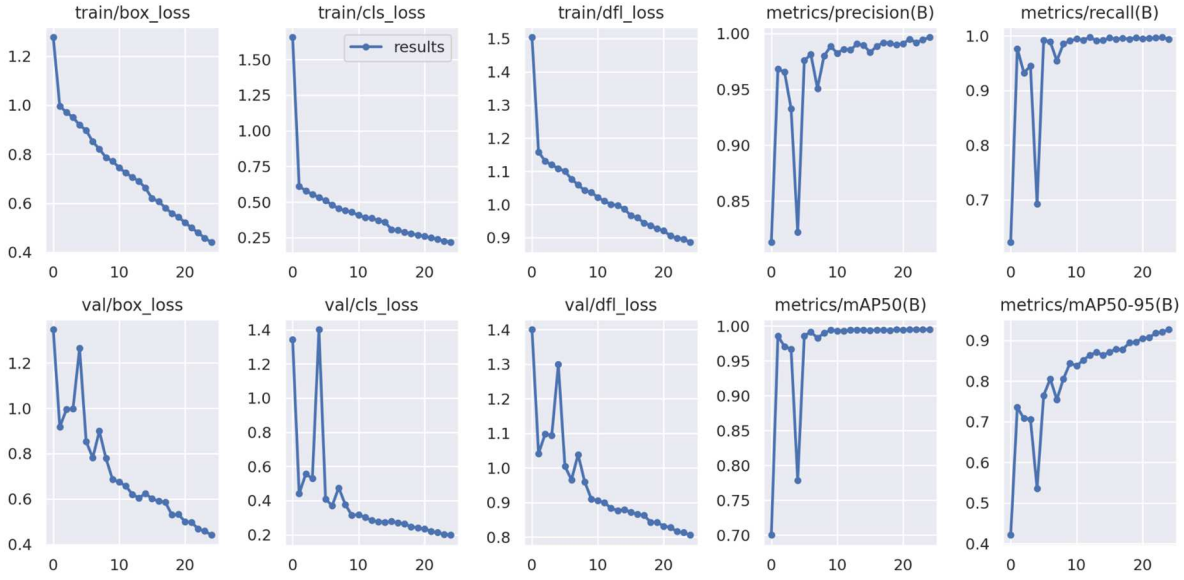


Fig. 6. YOLOv8 results without preprocessing

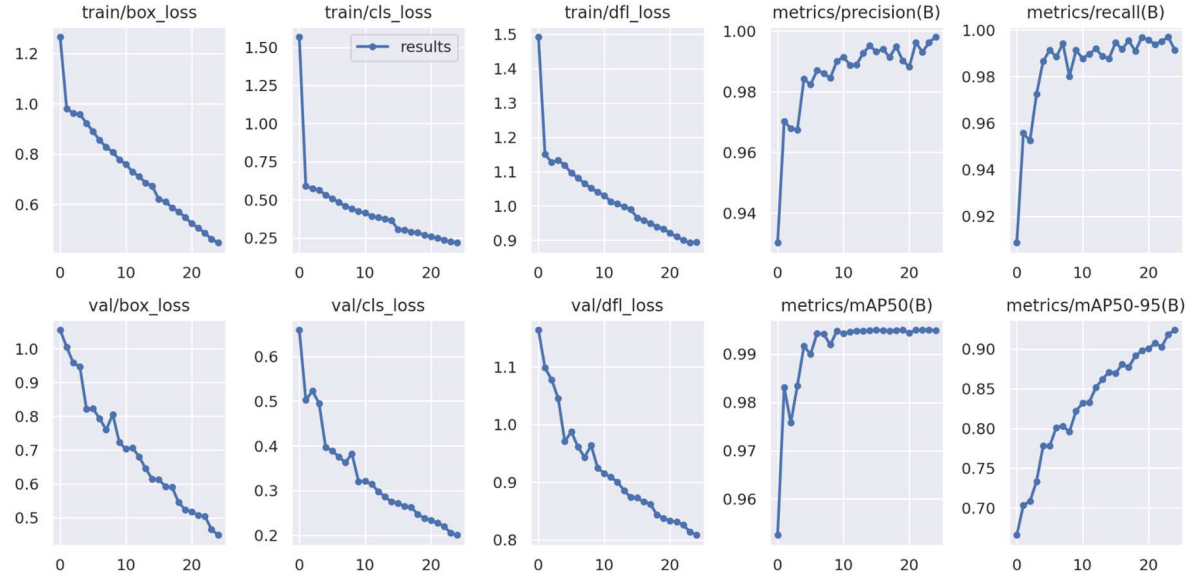


Fig. 7. YOLOv8 results with preprocessing technique 2: Dehazing

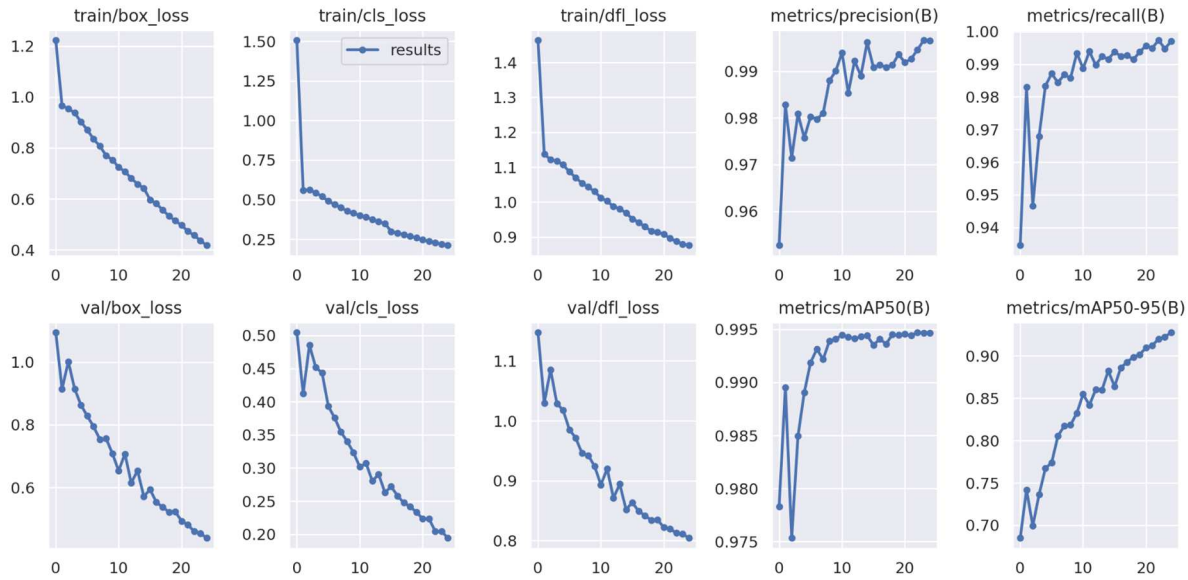


Fig. 8. YOLOv8 results with preprocessing technique 4: CLAHE.

Figures 6 to 8 show the results for each preprocessing technique. It shows the three types of loss functions of YOLOv8- box loss (how well the object is located inside the bounding box), classification loss (how the detected objects are classified), and DFL loss(for handling class imbalance) for training and validation. It also shows the metrics, including precision, recall, mAP50, and mAP50-95. Figure 6 shows the graph before preprocessing. Figure 7 shows the results of the dehazing technique. Figure 8 shows the results of the CLAHE technique. It can be inferred that the preprocessing techniques improve the overall performance of the model by showing a decrease with respect to loss functions and an increase with respect to output metrics. These techniques are compared based on the results shown in Table 3.

Table 3 shows that techniques 2(dehazing) and 4(CLAHE) have improved recall and mAP scores. So, either of these techniques can be applied in the preprocessing step before starting training to get improved results for handling unclear or turbid images.

4.2.2 Experimental Results comparing the YOLO models.

The YOLOv8 and the modified YOLOv8 are trained for 50 epochs. The object-detecting technology known as YOLO, an acronym for "You Only Look Once," has gained significant recognition for its exceptional efficiency and quickness. The YOLOv8 version, specifically, represents a more sophisticated iteration. In this study, the conventional YOLOv8 model and its modified version, called mYOLOv8, were trained on the Aquatic WeightNet dataset for a total of 50 epochs, and the metrics and losses are shown in Figures 9 and 10.

Table 4
Experimental Result for Aquatic WeightNet on different YOLOv8 architectures

Methodology	**Precision**	**Recall**	**mAP50**	**mAP 50-95**
YOLOv8	0.991	0.996	0.995	0.895
mYOLOv8	0.998	0.997	0.994	0.899

The training duration indicates that both models underwent 50 complete iterations of processing the full dataset. Table 4 offers valuable data regarding the performance of these models across a range of criteria.

The precision of YOLOv8, denoting the proportion of accurately predicted positive observations out of the total anticipated positives, is measured at 0.991. This finding indicates that around 99.1% of the model's detections were deemed to be accurate. In contrast, mYOLOv8 scored 0.998, indicating that approximately 99.8% of its detections were accurate. This suggests that the changed version exhibited a lower rate of false-positive detections than the standard version. Another important metric to consider is recall, which calculates the proportion of properly predicted positive observations in relation to all observations in the actual class. The YOLOv8 model achieved a high detection rate of around 99.6% for the identified items, as indicated by its score of 0.996. In contrast, the mYOLOv8 algorithm exhibited a detection rate of approximately 99.7% for all objects, representing a marginal enhancement compared to the conventional iteration.

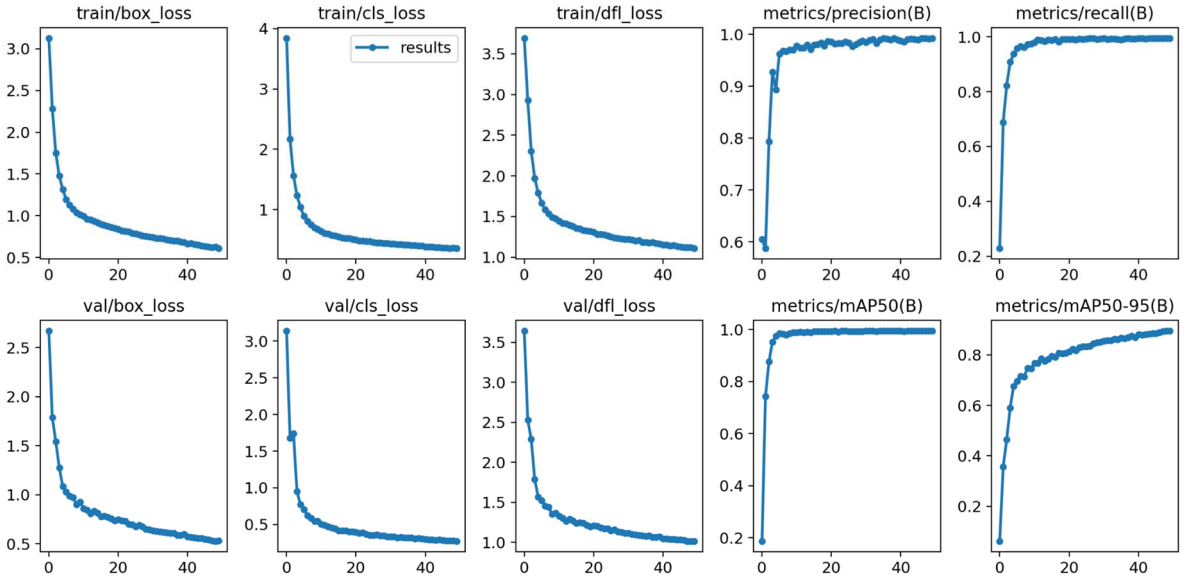


Fig. 9. YOLOv8 results for original YOLOv8

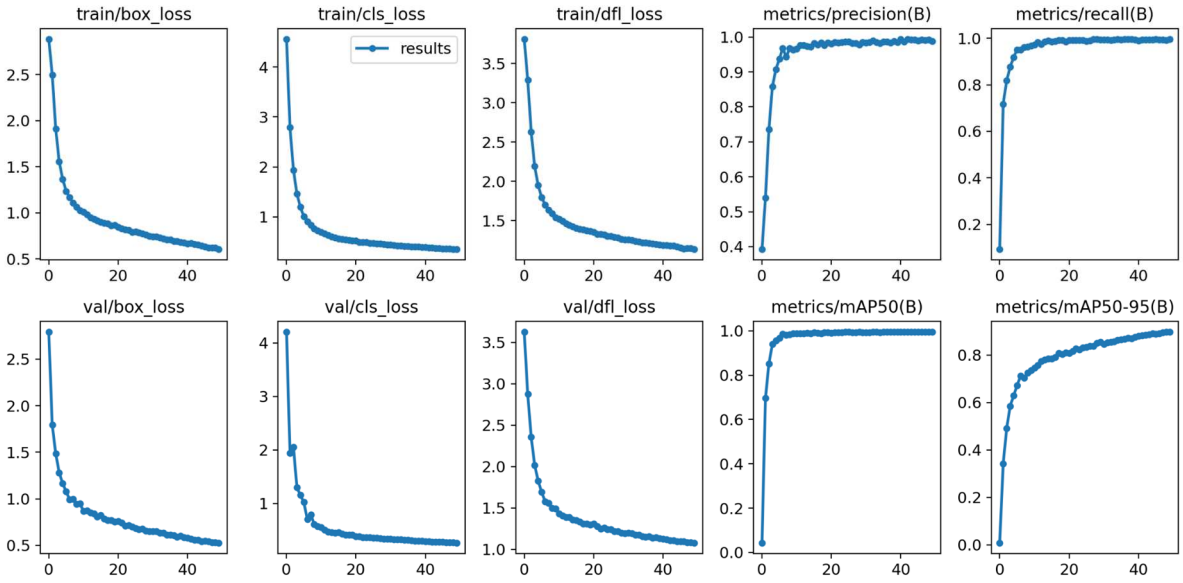


Fig. 10. YOLOv8 results for mYOLOv8.

The metric means Average Precision (mAP) at a 50% IoU threshold, also called mAP50, holds significant importance. The Intersection over Union (IoU) metric quantifies the degree of overlap between the predicted bounding box and the ground truth bounding box. The YOLOv8 model demonstrates a noteworthy mean average precision (mAP50) score of 0.995, which signifies a substantial level of accuracy in terms of average precision when considering a 50% intersection over union (IoU) threshold. The score achieved by the mYOLOv8 model, which is 0.994, demonstrates a slightly reduced but still outstanding level of performance. Finally, the mean average precision (mAP) metric is computed by considering various intersections over union (IoU) thresholds ranging from 50% to 95% and is denoted as mAP50-95. This statistic provides a thorough viewpoint on the performance of a model across different levels of overlap. The YOLOv8 model achieved a precision score of 0.895, indicating reliable performance even when more stringent overlap thresholds were used. When comparing the performance of several models, it was seen that the mYOLOv8 model attained a score of 0.899, indicating a slightly higher level of robustness across various thresholds.

The proposed method, which encompasses both the base YOLOv8 and its modified counterpart, demonstrates exceptional performance with mean Average Precision (mAP) scores of 0.995 and 0.994, respectively, achieved after 50 epochs. In a comparative context with similar studies, Tengtrairat et al., 2022 reported the following mAP scores after 300 epochs: Faster R-CNN (0.985), Mask R-CNN (0.991), RetinaNet (0.981), and YOLOv5 (0.903). The impressive performance of the proposed system can be attributed to the utilization of an advanced model—YOLOv8 and its modified version, mYOLOv8. Although the modified architecture does not exhibit improvement over the original model in terms of the mAP50 score, there is notable enhancement observed in the mAP50-95 scores when compared to the original model. This specific modification, involving the addition of detection heads and corresponding changes in the backbone, was implemented to accommodate different image sizes.

It is worth noting that while this modification proves effective in certain scenarios, it may not be universally optimal. One important point to note is that adding more detection heads will increase model complexity which may not be favourable in all circumstances. Alternative changes, such as adjusting the pooling function or incorporating different blocks into the backbone, could be explored depending on specific needs.

Future work could involve a more in-depth exploration of the YOLOv8 architecture, allowing for customization based on specific application requirements. This approach opens the door to further refinement and optimization of the model for diverse use cases.

In conclusion, it can be observed that both versions of YOLOv8 had a commendable performance on the Aquatic WeightNet dataset. However, it is worth noting that the mYOLOv8 variant exhibited a little advantage in terms of precision, recall, and mean average precision (mAP) at the 50-95 intersection over the union (IoU) threshold. The findings suggest that the architectural alterations made to YOLOv8 yielded positive outcomes, with only slight improvements. Even marginal enhancements might result in substantial practical consequences in situations characterized by high performance, rendering this insight attractive.
It can be inferred from Table 4 which shows the comparison of the models YOLOv8 and mYOLOv8 that there is an increase in precision and recall for the mYOLOv8 model. Though there is a slight increase in mAP50, mAP50-95 shows a significant increase.

4.2.3 Experimental Results for Linear Regression to estimate biomass.

Linear regression is used for each proposed model. Linear regression is a widely employed statistical technique that is foundational for modeling the association between a dependent variable and one or more independent variables. Within the framework of the study, linear regression techniques were employed for several tasks pertaining to the assessment of fish biomass. Table 5 provides an overview of the approaches employed and their respective measures. The Depth Estimator, which is designed to estimate the depth of fish, exhibits a Mean Absolute Error (MAE) of 0.370. The measure in question computes the mean absolute difference between projected and actual values, regardless of their positive or negative direction. Smaller values are more desirable, indicating that the model's predictions exhibit excellent proximity to the actual values.

Table 5
Linear Regression Results

Methodology	MAE	R^2	MAE/R^2
Depth Estimator	0.370	1.392	1.689
Pixel To Cm Convertor	0.892	0.849	1.294
Biomass Estimator	1	$1.97e^{-12}$	$2.73e^{-12}$

The Coefficient of Determination (R^2) is 1.392, which signifies the proportion of variance in the dependent variable that can be accurately predicted by the independent variable(s). The R^2 value is typically bounded within the range of 0 and 1. A score that is closer to 1 signifies a higher level of model fit. Nevertheless, in this scenario, a score exceeding 1 indicates the presence of inconsistencies or errors. The derived measure MAE/R^2 has a value of 1.689. Its meaning is contingent upon the specific context. However, it is generally assumed that a lower value would be more desirable. The objective of the Pixel to Centimeter Converter is to facilitate the conversion of pixel measurements inside a picture to corresponding centimeter values, hence enabling the determination of real-world dimensions. The model's Mean Absolute Error (MAE) is 0.892, indicating the average absolute difference between the predicted and actual values. The coefficient of determination (R^2) is 0.849, representing the proportion of the variance in the dependent variable that the independent variables can explain. The ratio of MAE to R^2 is calculated as 1.294, providing a measure of the relative magnitude of MAE in relation to the explained variance by the model.

The Biomass Estimator, a tool used to determine the true biomass of fish, exhibits a Mean Absolute Error (MAE) of 1. The coefficient of determination (R^2) for biomass estimation using linear regression is found to be approximately 0 at 1.97E-12, suggesting a negligible linear association between the anticipated and actual values. The calculated derived measure MAE/R^2 is equal to 2.73E-12. Upon examination of the obtained findings, it is evident that the performance of linear regression was suboptimal

for the given tasks. The obtained R^2 value of 0.37, which may be derived from an alternate table, indicates that the model can account for merely 37% of the dependent variable's variance, signifying a rather weak relationship. Furthermore, the elevated Mean Absolute Error (MAE) values, particularly shown in the Biomass Estimator, suggest significant inaccuracies in the predictive models. The R^2 coefficient of determination for the Biomass Estimator is around 0, suggesting a negligible linear relationship between the projected and observed values. This observation strongly suggests that linear regression is not an appropriate method for estimating biomass in this circumstance. Therefore, it is probable that alternative modeling methodologies need to be investigated in order to identify a more appropriate and efficient methodology.

The results suggest that linear regression did not perform well for these tasks. For instance, the R^2 value of 0.37 mentioned suggests that only 37% of the variability in the dependent variable can be explained by the model, which is not strong. Moreover, the high MAE values, especially for the Biomass Estimator, indicate considerable prediction errors. Furthermore, the R^2 value of the Biomass Estimator is almost 0, indicating that there is practically no linear correlation between the predicted and observed values. This is a vital sign that linear regression is unsuitable for biomass estimation in this context. The conclusion drawn from these observations is that while linear regression is a valuable tool in many scenarios, it might not be the best choice for this specific dataset or task. As observed from Table 5 which provides an overview of the approaches employed and their respective measures, R^2 is 0.37, which is not an excellent value, and the MAE and RMSE of 1.392 and 1.689 are also high. This indicates that linear regression did not give satisfactory results. Therefore, there is a need to find the best-performing algorithm for the dataset.

4.2.4 Experimental Results for finding the best depth estimator.

The findings for various regression models, which seek to estimate depth, given the length and width in pixels, are presented in Table 6. The performance of these models has been evaluated using many metrics, such as training time, prediction time, variance (also referred to as (R^2)), Mean Absolute Error (MAE), and the MAE-to-(R^2) ratio. Commencing with the Linear Regression Analysis, the linear regression model's coefficient of determination (R^2) is calculated to be 0.37. The metric in question quantifies the extent to which the variability of the dependent variable (namely, depth) can be accounted for by the independent variables within a regression model. A score of 1 signifies optimal prediction, whereas a value of 0 shows that the model does not enhance prediction beyond the use of the mean of the dependent variable. As indicated by its R^2 value of 0.37, the linear regression model accounts for merely 37% of the variability observed in the dependent variable. This finding implies that alternative models may be more suitable for accurately representing the dataset under consideration.

The performance of the K Neighbours Regressor closely approximates that of the model, as evidenced by an R^2 value of 0.796. Nevertheless, the Extra Trees Regressor demonstrates superior performance compared to other models, as evidenced by its R^2 value of 0.870, indicating that it accounts for 87% of the variability in the dependent variable. Additionally, it exhibits the lowest Mean Absolute Error (MAE) of 0.637. The Random Forest Regressor exhibits a strong correlation with an R^2 value of 0.826. In contrast, the Decision Tree Regressor exhibits lower performance, as seen by its R^2 value of 0.644. In the context of linear regression, it can be argued that the Lasso and Ridge models, which are regularization approaches, do not provide substantial enhancements compared to the conventional linear regression approach. Furthermore, considering the computational aspect, it can be observed that the Random Forest Regressor exhibits the longest training time. However, the potential improvement in performance may warrant the investment of additional effort. performance of Linear Regression, Lasso, and Ridge models is deemed unsatisfactory despite their efficient training times.

In conclusion, the Linear Regression model exhibited suboptimal performance, whereas the Extra Trees and Random Forest regressions showed superior suitability for the given dataset and the job of depth prediction. However, optimal model selection is contingent upon practical considerations such as processing resources, forecast speed, and the required level of accuracy. Table 6 shows that MAE/R^2 is the lowest for the Extra Trees Regressor. Hence, this is selected as the regression technique, and the optimal parameters for this regressor are found- the number of estimators is 100, and the criterion is the squared error. MAE of 0.638 and R^2 score of 0.864 are observed from Figure 11 thus signifying how well the independent variable can explain the variance in the dependent variable.

Table 6
Finding the best model for depth estimator

Model	Training time	Prediction time	Variance	MAE	R^2	MAE/ R^2
Gradient Boosting Regressor	0.681s	0.003s	0.803	0.645	0.803	0.803
K Neighbors Regressor	0.007s	0.005s	0.796	0.648	0.796	0.813
Extra Trees Regressor	1.097s	0.065s	0.870	0.637	0.870	0.733
Random Forest Regressor	2.518s	0.052s	0.826	0.646	0.826	0.783
Decision Tree Regressor	0.047s	0.002s	0.644	0.762	0.644	1.184
Linear Regression	0.006s	0.001s	0.325	1.480	0.324	4.565
Lasso	0.003s	0.002s	0.325	1.481	0.324	4.570
Ridge	0.004s	0.002s	0.325	1.48	0.324	4.565
Extra Trees Regressor	1.097s	0.065s	0.870	0.637	0.870	0.733

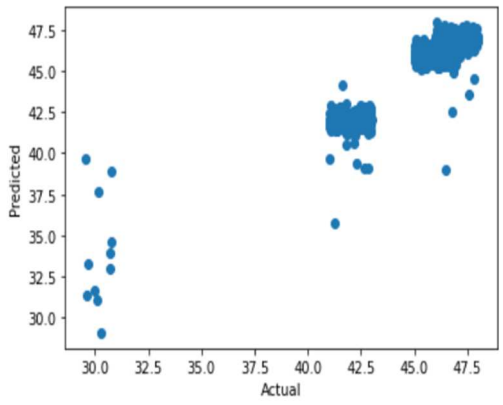


Fig. 11. Actual v/s predicted for depth estimation.

4.2.5 Experimental Results for finding the best pixel to cm converter.

This section aims to identify the most effective and precise model for converting pixel data to centimeters (cm) given the length (pix) , width (pix), and depth. This is of utmost importance in accurately estimating biomass in aquaculture settings. The performance of different regression models for the conversion job is presented in Table 7. The model known as the K Neighbours Regressor used 0.021 seconds for the training process and 0.013 seconds for making predictions. The obtained variance score of 0.998 indicates a robust correlation between the observed and projected values. The Mean Absolute Error (MAE) value of 0.022 signifies the average discrepancy between the model's predictions and the actual values, measured in centimeters. The model demonstrates a high level of explanatory power, as evidenced by an R^2 value of 0.998, indicating that it can account for 99.8% of the variability observed in the data. The Mean Absolute Error (MAE) ratio to the coefficient of determination (R^2) is 0.022, providing insight into the relationship between the magnitude of the error and the amount of variance explained. The Extra Trees Regressor exhibited a training time of 0.721 seconds and a prediction time of 0.125 seconds. The model attained a variance score and R^2 score of 1, suggesting that it accurately predicted the observed data with a high degree of precision. Similarly, the model's Mean Absolute Error (MAE) was determined to be 0, indicating the absence of any prediction mistakes. The Random Forest Regressor exhibited a more significant training duration of 3.271 seconds and a predicted duration of 0.029 seconds. The performance measures exhibited by the system were highly accurate, as indicated by a variance of 0.999 and an R^2 score of 0.999. The model achieved a minimum Mean Absolute Error (MAE) of 0.014. The Decision Tree Regressor demonstrated a notably short training time of 0.065 seconds and an even faster prediction time of 0.003 seconds. Like the extra trees regressor, the decision tree regressor demonstrated exceptional performance in terms of variance, R^2 score, and MAE, highlighting its resilience and suitability for the given job. The Linear Regression model was trained in 0.013 seconds and made predictions in 0.003 seconds.

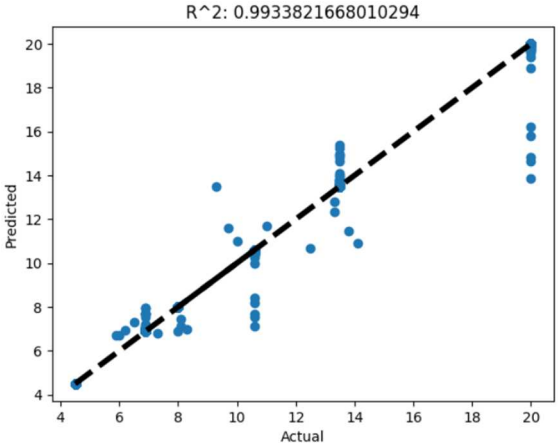


Fig. 12. Actual v/s predicted for pixel to cm conversion.

Nevertheless, the performance parameters of this particular model exhibited distinct differences compared to the other models. It displayed a variance and R^2 score of 0.899, along with a somewhat elevated mean absolute error (MAE) of 0.912. The ratio between the Mean Absolute Error (MAE) and the coefficient of determination (R^2) was found to be 1.015. The Lasso regression model used 0.012 seconds for training and 0.004 seconds for making predictions. The performance indicators of the model in question show a slight decrease compared to those of the Linear Regression model. Specifically, the model's variance and R^2 score were measured at 0.892, while the mean absolute error (MAE) was found to be 0.976. The ratio between the Mean Absolute Error (MAE) and the coefficient of determination (R^2) was recorded as

Table 7
Finding the best model for pixel to cm converter

Model	Training time	Prediction time	Variance	MAE	R^2	MAE/R^2
K Neighbors Regressor	0.021s	0.013s	0.998	0.022	0.998	0.022
Extra Trees Regressor	0.721s	0.125s	1	0	1	0
Random Forest Regressor	3.271s	0.029s	0.999	0.014	0.999	0.014
Decision Tree Regressor	0.065s	0.003s	1	0	1	0
Linear Regression	0.013s	0.003s	0.899	0.912	0.899	1.015
Lasso	0.012s	0.004s	0.892	0.976	0.892	1.094
Ridge	0.010s	0.004s	0.899	0.912	0.899	1.015
K Neighbors Regressor	0.021s	0.013s	0.998	0.022	0.998	0.022
Extra Trees Regressor	0.721s	0.125s	1	0	1	0

1.094. Finally, the Ridge regression model exhibited performance metrics comparable to those of the Linear Regression model. The model underwent training for 0.010 seconds, followed by prediction in 0.004 seconds. Furthermore, its variance, R^2 score, and mean absolute error (MAE) exhibited similarity to those of the Linear Regression model.

In conclusion, certain models, including Extra Trees and Decision Tree Regressors, attained flawless scores in the context of the pixel-to-cm conversion work. Conversely, Linear Regression, Lasso, and Ridge models exhibited suboptimal results, thereby underscoring the diverse performance of regression models in this particular assignment.

Table 7 shows that MAE/R^2 is the lowest for the Random Forest Regressor. Hence, this is selected as the regression technique, and the optimal parameters for this regressor are found: the number of estimators is 100, and the criterion is absolute error. MAE of 0.014 and R^2 of 0.99 can be observed from Figure 12.

4.2.6 Sensitivity Analysis

Sensitivity analysis is a technique that helps understand how the machine learning model responds to changes in the inputs or parameters. Sensitivity analysis was conducted on YOLOv8 and the modified version mYOLOv8, and the results were obtained by varying various input parameters. Similarly, a scatter plot is obtained that explains how biomass depends on the length and width of the fish.

The graphs show how different inputs affect the output metrics. All parameters are kept the same except for the one that was studied. For example, Figure 13 shows how the output values of precision, recall, mAP50 and mAP50-95 are affected when the number of epochs is changed from 10 to 50. Increasing

the number of epochs from 10 to 50 has a greater effect on the output metrics in that the performance increases as the number of epochs increases. For the input parameter cos_lr (cosine learning rate scheduler), as shown in Figure 14, using a cosine curve for the learning rate has different effects on the output. For some output metrics, not using the cosine curve is better, while for others, using the curve is better. For initial and final learning rates, shown in Figure 15 and Figure 16, respectively, the outputs differ, but it can be observed that it is good to stop at the optimal level. When increasing the learning rates to 0.02, the models perform worse. With respect to optimizers, it can be seen from Figure 17 that RMSProp performs worse than SGD or AdamW for both YOLOv8 and mYOLOv8. Thus, it can be inferred that the choice of optimizer affects the model's performance. For momentum, too, the values vary, and higher values do not indicate higher performance, as indicated by Figure 18. Thus, it can be concluded that the optimal parameters need to be chosen with rigorous analysis as they directly affect the performance of the models.

The scatter plot in Figure 19 shows the relationship between length and width v/s biomass. For more length and more width, the biomass is higher. Although it is not strictly linear, the relationship between length, width and biomass can be observed.

4.2.7 Experimental Results for finding the best biomass estimator.

The experimental findings presented in this section primarily center around identifying the optimal model for estimating fish biomass given the turbidity, length, and width in pixels, and centimeters. A range of regression models were evaluated, and the resulting performance measures and the corresponding computation durations were documented. The Gradient Boosting Regressor model required 1.484 seconds for the training process and a minimal 0.005 seconds for generating predictions.

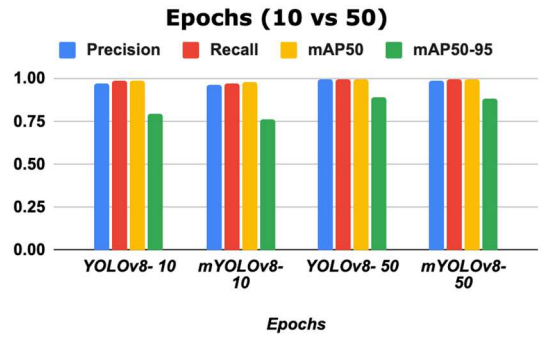


Fig. 13. Epochs 10 vs. 50

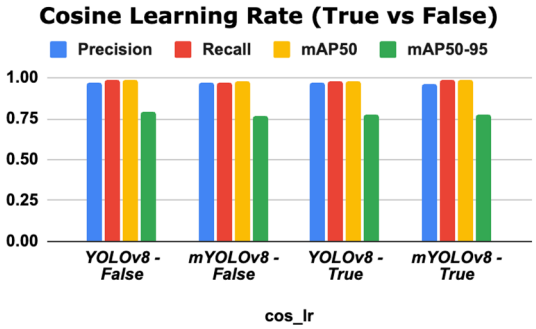


Fig. 14. cos_lr False vs. True

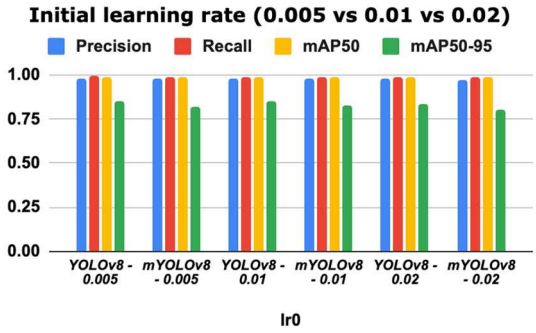


Fig. 15. Initial learning rate 0.005 vs.. 0.01 vs. 0.02

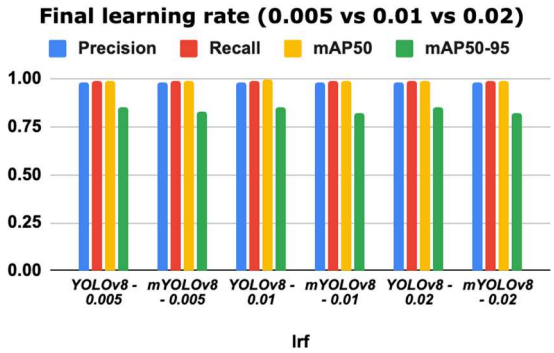


Fig. 16. Final learning rate 0.005 vs 0.01 vs 0.02

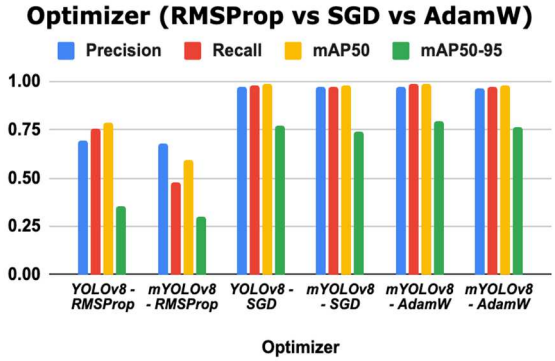


Fig. 17. Optimizer RMSProp vs SGD vs AdamW

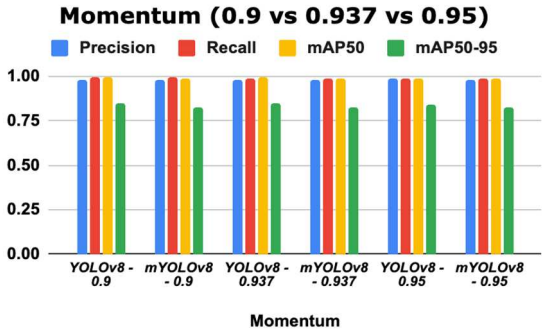


Fig. 18. Momentum 0.9 vs 0.937 vs 0.95

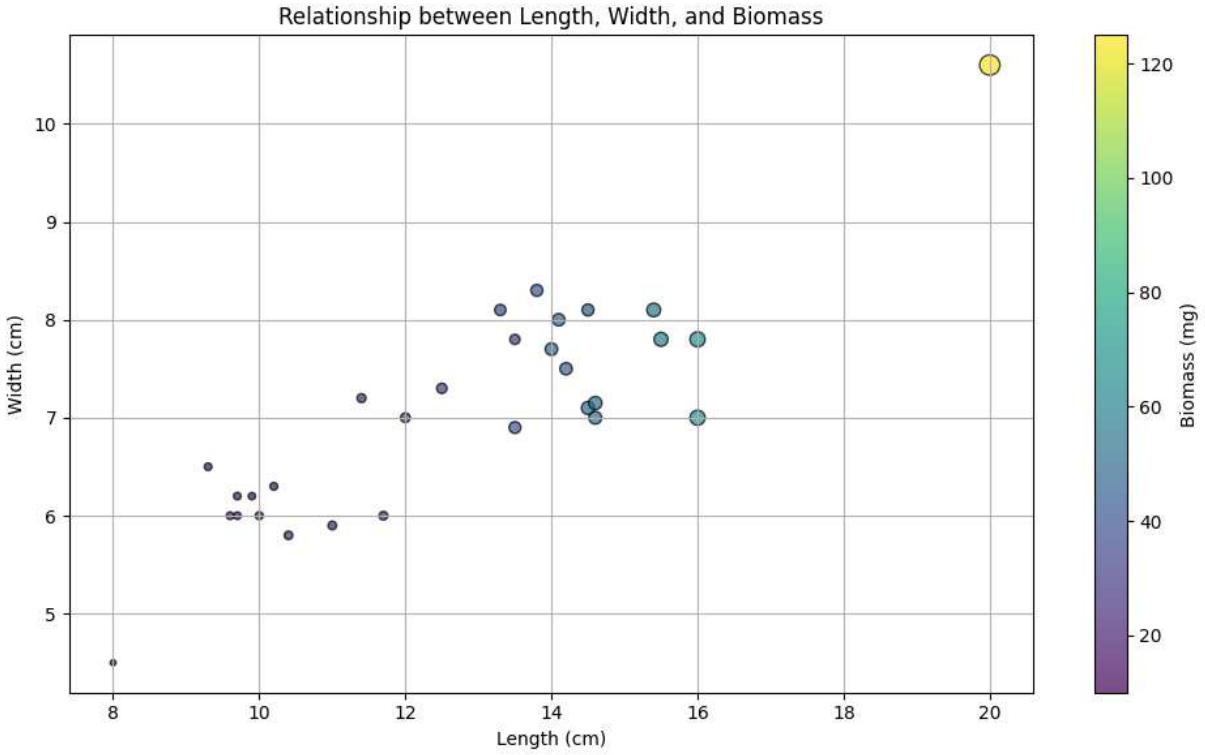


Fig. 19. Scatter plot showing relationship between length, width and biomass

The model demonstrated a significantly high variance score and R^2 score, both measuring at 0.999. This indicates that the model could account for 99.9% of the variability observed in the response data. The Mean Absolute Error (MAE) exhibited a value of 0.013. Upon evaluating the MAE in relation to the R^2 score, the resulting ratio was also found to be 0.013. The K Neighbours Regressor model exhibits notable efficiency in the training phase, with a training time of merely 0.028 seconds. However, its prediction time is slightly higher, at 0.024 seconds. Additionally, the model demonstrated exceptional performance, as evidenced by a variance score and a R^2 score of 0.998. The MAE of the model had a substantially higher value of 0.145, resulting in a MAE/R^2 ratio of 0.146. The Extra Trees Regressor model demonstrated a training time of 0.489 seconds and a prediction time of 0.050 seconds. The observed high variance and R^2 value of 0.999 in the previous models were replicated. The mentioned models' Mean Absolute Error (MAE) was remarkably low, measuring at 0.004. Consequently, this yielded an MAE/R^2 ratio of 0.004. The Random Forest Regressor model exhibited a training time of 0.895 seconds, while the prediction time was recorded as 0.013 seconds. Once more, the variance and R^2 scores exhibited a value of 0.999. The model's mean absolute error (MAE) was determined to be 0.010, accompanied by an MAE/R^2 ratio of 0.010. The Decision Tree Regressor model demonstrated exceptional efficiency in terms of both training (0.012 seconds) and prediction (0.001 seconds) durations. The variance and R^2 scores of the current models were consistent with the previous models, both measuring at 0.999. The model's mean absolute error (MAE) was determined to be 0.009, leading to a ratio of MAE to R^2 of 0.009. The linear regression model had a notably swift training time of 0.005 seconds and a similarly rapid prediction time of 0.001 seconds. Nevertheless, the system's performance measures exhibited a minor decline, as indicated by a variance and R^2 score of 0.993. This model's Mean Absolute Error (MAE) had a

significantly higher value of 1.628, resulting in an MAE/R^2 ratio of 1.640. The Lasso model exhibited a training time of 0.035 seconds and a prediction time of 0.002 seconds. The data exhibited a variance score 0.988 and a R^2 value 0.988. The Mean Absolute Error (MAE) exhibited the highest value of 4.131 across all the models, leading to an MAE/R^2 ratio of 4.181. The training and prediction times for Ridge Regression were observed to be 0.005 seconds and 0.002 seconds, respectively. The model obtained a variance and R^2 score of 0.993. The model's Mean Absolute Error (MAE) was calculated to be 1.648. Additionally, the MAE to R-squared (MAE/R^2) ratio was determined to be 1.660.

In brief, whereas most models exhibited nearly impeccable variance and R^2 scores, notable disparities were seen in their mean absolute errors (MAEs). The Extra Trees Regressor had the lowest Mean Absolute Error (MAE), whereas the Lasso model demonstrated the greatest MAE. The training and prediction timeframes exhibited variability among the models, providing valuable insights into their computational efficiency.

Table 8 shows the results for the various regression techniques to find the best one for the biomass estimator. It is observed that MAE/R^2 is the lowest for the extra trees regressor. MAE is 0.004 and R2 is 0.99. The main emphasis of these findings is around three key metrics, namely Mean Absolute Error (MAE), R-squared (R^2), and the ratio of MAE to R^2 (MAE/R^2). The "Depth Estimator" methodology yields a Mean Absolute Error (MAE) value of 0.6372. The Mean Absolute Error (MAE) metric assesses the average magnitude of errors between expected and observed values. It indicates the proximity of forecasts to the actual results. The mean absolute error (MAE) of 0.6372 indicates that, on average, the depth estimator's predictions exhibit a deviation from the actual values by around 0.6372 units. The coefficient of determination (R^2) for the Depth Estimator is 0.8695. The coefficient of determination, denoted as R^2, quantifies the fraction of the variance in the dependent variable that can be accounted for by the independent variables included in a regression model. A coefficient of determination (R^2) equal to 0.8695 indicates that the model can explain about 86.95% of the variance in the data. This relatively high value implies a strong correspondence between the model and the observed data, indicating a favorable fit level. The Mean Absolute Error (MAE) ratio to the coefficient of determination (R^2) for the Depth Estimator is computed as 0.7328. The ratio offers a modified assessment of the discrepancy in relation to the variance that can be accounted for. A smaller ratio often signifies superior model performance when considering both error and explained variance.

Regarding the "Pixel to Cm Converter," the Mean Absolute Error (MAE) is calculated to be 0.0141. This value signifies that, on average, the predicted values differ from the actual outcomes by a mere 0.0141 units. Consequently, it can be inferred that the converter exhibits a notable degree of precision. The R-squared coefficient has a noteworthy value of 0.9991, signifying that the model effectively accounts for about 99.91% of the observed variability in the dataset. This outcome suggests a remarkably strong alignment between the model and the data. The ratio between the mean absolute error (MAE) and the coefficient of determination (R^2) for this methodology is equivalent to the MAE itself, which is 0.0141.

In conclusion, the "Biomass Estimator" approach demonstrates a Mean Absolute Error (MAE) of 0.0041, indicating a remarkably high degree of predictive accuracy and a low departure from the observed results. The coefficient of determination (R^2) is 0.9999, indicating that the model effectively accounts for nearly all (99.99%) of the variance seen in the dataset. The ratio between the mean absolute error (MAE) and the coefficient of determination (R^2) for this estimator is 0.0041, which is consistent with the value of the MAE.

Table 8

Finding the best model for the biomass estimator.

Model	Training time	Prediction time	Variance	MAE	R^2	MAE/R^2
Gradient Boosting Regressor	1.484s	0.005s	0.999	0.013	0.999	0.013
K Neighbors Regressor	0.028s	0.024s	0.998	0.145	0.998	0.146
Extra Trees Regressor	0.489s	0.050s	0.999	0.004	0.999	0.004
Random Forest Regressor	0.895s	0.013s	0.999	0.010	0.999	0.010
Decision Tree Regressor	0.012s	0.001s	0.999	0.009	0.999	0.009
Linear Regression	0.005s	0.001s	0.993	1.628	0.993	1.640
Lasso	0.035s	0.002s	0.988	4.131	0.988	4.181
Ridge	0.005s	0.002s	0.993	1.648	0.993	1.660
Gradient Boosting Regressor	1.484s	0.005s	0.999	0.013	0.999	0.013

Moreover, all three approaches exhibit high R^2 values, which suggest robust model fits. The consistently low mean absolute error (MAE) values observed indicate that the approaches employed exhibit a high level of accuracy in their predictive capabilities.

Table 9
Optimized Regression Result

Methodology	**MAE**	**R^2**	**MAE/R^2**
Depth Estimator	0.6372	0.8695	0.7328
Pixel To Cm Convertor	0.0141	0.9991	0.0141
Biomass Estimator	0.0041	0.9999	0.0041

The MAE/R^2 ratios offer a comprehensive assessment of the models' capabilities. Observing low values across the various techniques suggests that the models are effectively capturing the data patterns while minimizing errors. Grid search is applied at each step to find the best-performing regression technique, and the results are tabulated, as shown in Table 9. Extra trees regression performs well for the depth estimator, giving an MAE of 0.6372 and R^2 of 0.869. For the pixel-to-cm converter, random forest performs best with MAE of 0.014 and R^2 of 0.999. For the biomass estimator, extra trees perform best by giving an MAE of 0.004 and R^2 of 0.999.

4.2.8 Comparison of fish biomass estimation models with and without depth.

The models are compared with and without depth, and the results yielded are as follows. This study uses several models to examine the comparability of fish biomass estimation, focusing on the impact of integrating depth into the estimation procedure (as shown in Table 10). This study assesses two techniques, specifically P2CM and BE, by examining their performance in terms of Mean Absolute Error (MAE) and R-squared (R^2) metrics. In the context of the P2CM methodology, the Mean Absolute Error (MAE) is calculated as 0.044, indicating the average error size in the biomass estimates. Additionally, the R-squared (R^2) value is reported to be 0.992. This finding indicates that a

substantial majority (99.2%) of the variability in the observed data may be accurately estimated using the proposed model. Nevertheless, including depth in the P2CM model leads to a noteworthy enhancement in accuracy. The mean absolute error (MAE) decreases to 0.015, indicating a decrease in the average error.

Table 10
Comparison of models with and without depth

Methodology	**MAE**	**R^2**
Without depth- P2CM	0.044	0.992
With depth- P2CM	0.015	0.998
Without depth- BE	1.694	0.991
With depth- BE	0.003	0.999

Additionally, the coefficient of determination (R^2) increases to 0.998, suggesting a higher level of predictability of variance (99.8%) from the model. When using the BE approach, it is observed that the MAE significantly increases to 1.694, suggesting a more significant average error. However, the R^2 value of the model remains impressive, standing at 0.991. This indicates that the model can accurately predict around 99.1% of the observed variance. Remarkably, incorporating depth into the BE model yields a significant enhancement. The mean absolute error (MAE) experiences a significant decrease to 0.003, suggesting an almost insignificant average error in estimation. Concurrently, the coefficient of determination (R^2) attains an almost perfect value of 0.999, indicating a high level of prediction, approximately 99.9%, regarding the variability explained by the model.

It can be inferred that including depth in the biomass estimating process is advantageous for both P2CM and BE techniques. This inclusion improves accuracy, as seen by the observed decrease in mean absolute error (MAE) values. Furthermore, the models consistently exhibit strong R^2 values across various circumstances, with a minor increase observed when considering the depth.

Table 10 shows the results of the pixel-to-centimeter converter and Biomass Estimator with and without including depth as a factor. P2CM is the pixel-to-centimeter converter. BE is the biomass estimator. The MAE for the biomass estimator is 1.694, and it reduces to 0.003 when depth is included. R^2 increases from 0.991 to 0.999. Therefore, the depth information and, thus, the depth estimation model are crucial for estimating biomass.

At each step, the most suitable regression model is found. The last step outputs the estimated fish biomass. This is one of the many techniques that can be used to estimate fish biomass.

While the series of methods employed in this research has demonstrated commendable results, it is essential to acknowledge certain limitations that may impact the robustness and generalizability of the findings. Specifically, the model's performance may be compromised under extreme conditions such as low light, extreme salinity, or instances of significant fish overlap. These conditions represent challenges inherent in unpredictable and dynamic aquatic environments.

The implications of these limitations are crucial to consider when interpreting the results. In scenarios where the environmental conditions deviate substantially from the norm, fish detection and biomass estimation accuracy may be compromised.

Another limitation that needs to be addressed is the dataset. The current dataset is that of GIFT Tilapia. It is not very diverse. Therefore, applying the current model on extremely diverse datasets may not yield very accurate results, although training the current system on different fish species is feasible.

These limitations may affect the model precision, recall, and mAP values. Therefore, it is imperative to exercise caution when extrapolating the system's capabilities to scenarios significantly deviating from the conditions encountered during its development.

Fish biomass estimation is critical for sustainable fishery management. Policy recommendations based on this research are as follows:

- Encourage aquaculture companies to implement sustainable feeding techniques based on fish biomass estimates. Implement policies that encourage optimal feeding, reduce waste and environmental effect, and ensure the health and growth of fish populations.

- Combine fish biomass data with more comprehensive environmental monitoring systems. This technique will provide a comprehensive understanding of the impact of feeding practices on water quality, allowing for more sustainable aquaculture and reducing environmental deterioration.

- Provide research grants and funds to support studies to improve fish biomass estimating methodology. Government involvement will promote new research, resulting in more accurate and efficient approaches.

- Initiate comprehensive aquaculture training programs emphasizing improved fish biomass estimation methodologies. This will enable practitioners to collect reliable data independently, resulting in more informed stocking, feeding, and harvesting decisions.

Table 11
Comparison of results between the proposed methodology and the results obtained by Tengtrairat et al., 2022.

Technique	**Error in Weight Estimation**
Tengtrairat et al., 2022	±29g
Proposed Methodology	±2g

Table 11 compares the results between the proposed methodology and Naruephorn Tengtrairat et al., 2022. The proposed model has seen improvements in biomass estimation by achieving an error of ±2g against ±29g obtained by existing work.

There are potential sources of error in the research. Right from data collection, the possible sources of error need to be kept in mind. It is possible that there would be an error in measurement since the dataset was collected manually. Moreover, since the dataset used in this study is primarily GIFT Tilapia, the dataset is not diverse, and the model might not predict very well for every fish species. The preprocessing techniques used in this research may falter under extreme conditions, compromising biomass estimation accuracy. The sensitivity of selected regression models and detection algorithms to environmental factors and fish behavior can also contribute to errors if they are not adequately accounted for.

Several strategies can be employed to improve the underlying assumptions and mitigate potential sources of error. Enhancing dataset diversity by incorporating data from a broader range of fish species and environmental conditions can enhance the model's generalizability. Research into more robust preprocessing techniques resilient to extreme environmental conditions is imperative to ensure consistent and accurate image data for biomass estimation. Rigorous validation and sensitivity analyses of regression models and detection algorithms can identify potential errors and improve model robustness.

Potential inaccuracies might have a substantial impact on analysis outcomes. Inaccuracies caused by measurement errors, low dataset diversity, and the sensitivity of models and algorithms to environmental conditions and fish behaviour can result in inaccurate fish biomass estimates in turbid environments. These mistakes can weaken the credibility of study findings in the eyes of policymakers and the scientific community. Inaccurate fish biomass estimates can mislead policymakers and scientists, potentially leading to poor policy decisions and inefficient tactics for combating climate change and its effects on aquatic ecosystems. If not addressed and presented properly, the “wrongness” in the research findings

could result in misguided policy actions, insufficient resource allocation, and inefficient conservation initiatives. As a result, policymakers and the scientific community must be upfront about potential study errors and limitations in order to make well-informed decisions and design viable policies to fulfill climate change targets and preserve the sustainability of aquatic ecosystems.

The following reference papers draw parallels to compare and contrast with the existing research and point out the key findings. The papers used are: Tengtrairat et al., 2022, and Abinaya et al., 2022.
Each study focuses on addressing the challenges of estimating fish biomass in different environments, emphasizing non-invasive methods and utilizing innovative techniques. All the studies emphasize the integration of advanced technologies such as deep learning, computer vision, and artificial intelligence for fish biomass estimation, showcasing a global trend toward technological innovation in aquaculture.
Abinaya et al., 2022., and this study highlights the importance of non-invasive approaches to fish biomass estimation, aiming to optimize aquaculture efficiency and minimize harm to fish, reflecting a global focus on ecological sustainability. The studies emphasize the importance of precision and efficiency in fish biomass estimation, utilizing regression models and advanced detection algorithms for accurate calculations, indicating a global commitment to optimizing aquaculture practices.
In summary, the parallels drawn from these studies and this study demonstrate a global commitment to leveraging advanced technologies, developing non-invasive methodologies, and optimizing fish biomass estimation for aquaculture practices' ecological sustainability and efficiency. These parallels reflect a collective effort towards innovation and environmental responsibility in the aquaculture industry across different geographical regions.

5. Conclusion and Future Work

In this research, a system has been developed to estimate the biomass of fish in a turbid environment from underwater videos. The images are processed using CLAHE, dehazing, denoising, and color balance techniques, of which dehazing and CLAHE give the best results, with precision in CLAHE increasing by 2.04% and mAP50-95 in CLAHE improving by 3.7%. A modification in the YOLOv8 architecture is proposed, for which a recall of 0.997 and mAP50-95 of 0.899 is achieved, improving the revised YOLOv8 (mYOLOv8) to achieve an accuracy of 99.98%. The best-performing regression technique was selected using grid search and applied for depth estimation, pixel-to-centimeter conversion, and biomass estimation. Biomass is predicted with an MAE of 0.01 and an R^2 score of 0.99. However, it is imperative to recognize the limitations identified during this study, particularly the model's sensitivity to extreme environmental conditions such as low light, extreme salinity, and instances of fish overlap. These limitations emphasize the need for further research to refine the system's adaptability to a broader range of underwater scenarios. Future research should improve preprocessing techniques, incorporate deep learning advancements, and tailor datasets to address specific challenges such as extreme occlusion. Exploring modifications to other components of the YOLOv8 architecture and understanding the impact of environmental factors on fish biomass will be crucial for enhancing the model's performance under varying conditions. This research work highlights the relationship between the turbidity of water and biomass. Additionally, this research can help learn the feeding pattern of fish by recommending the feed amount. The study's key findings are as follows: The use of image processing techniques, specifically Contrast Limited Adaptive Histogram Equalization (CLAHE) and dehazing, resulted in considerable increases in precision and mean average precision (mAP) for fish biomass estimation. Furthermore, a modified YOLOv8 architecture, known as mYOLOv8, was suggested and demonstrated outstanding accuracy in fish detection. The best regression technique allowed for precise and very accurate biomass estimation.

Regarding engineering design, the proposed image processing algorithms and modified YOLOv8 architecture provide useful insights for designing underwater monitoring systems, helping to advance the development of innovative technologies for aquatic environment evaluation. Furthermore, accurate fish biomass estimates affect aquaculture regulation and policy design. The study's findings can

help shape policies targeted at promoting sustainable fish farming practices and ecosystem conservation. Regarding energy systems, accurate assessment of fish biomass helps optimize energy-efficient aquaculture operations, which aligns with the aquaculture industry's commitment to sustainable energy resource management. Furthermore, the study's emphasis on ESG (environmental, social, and governance considerations) is consistent with the increasing importance of sustainable and socially responsible operations. Aquaculture operations that accurately determine the fish population in aquatic environments can better control resource utilization, feed conversion ratios, and waste output, resulting in a lower environmental effect. Ethical and socially responsible aquaculture operations include ensuring the health of aquatic environments and local communities.

For future enhancements of the current research on fish biomass estimation using the "Aquatic WeightNet" dataset, a comprehensive approach integrating both novel and refined methods is proposed. This includes the adaptation of the Genghis Khan Shark Optimizer (Hu et al., 2023) to improve the segmentation and tracking of fish in turbid underwater conditions. By mimicking the efficient hunting behavior of sharks, this optimizer can refine image preprocessing techniques like dehazing and CLAHE, enhancing image clarity and aiding the YOLOv8 model in more accurate fish detection and biomass estimation. Development of the Geyser-Inspired Algorithm (Hu et al., 2024) will focus on dynamically adjusting image preprocessing parameters to respond to variations in underwater visibility, leveraging eruptive behavior modeling to enhance the real-time performance of image preprocessing techniques, ensuring consistent image quality and improved accuracy in fish size and biomass predictions.

The Prairie Dog Optimization Algorithm (Ezugwu et al., 2022) will optimize collaboration among multiple detection models, improving the accuracy and efficiency of identifying fish characteristics, which could be particularly beneficial for the integration and performance enhancement of various regression models used for estimating fish dimensions and biomass. The Dwarf Mongoose Optimization Algorithm (Agushaka et al., 2022) will foster enhanced collaboration among different preprocessing and detection models, facilitating more effective exploitation of the dataset and improving the precision of fish size and biomass measurements, crucial for optimizing feed quantities and enhancing aquaculture efficiency. The Gazelle Optimization Algorithm (Agushaka et al., 2023) will enhance the dynamic response of detection models to varying turbidity levels in underwater environments by dynamically adjusting the detection and regression models to evade suboptimal solutions and converge more rapidly towards optimal settings for biomass estimation.

Furthermore, the introduction of deep learning techniques to preprocessing under extreme conditions such as low light, extreme occlusion, or overlapping of fish can address unique challenges by tailoring datasets to specific use cases and training deep learning models on these datasets. Incorporating Bidirectional LSTM with GloVe (Abualigah et al., 2024, Chen et al., 2020) embeddings could enhance the detection and classification of fish species within the dataset, capturing contextual and semantic information that improves the accuracy of fish biomass estimates. Continuing to modify the architecture of deep learning models like YOLO, to better suit specific requirements, includes altering various parts of the architecture as needed to handle different environmental factors, such as salinity, which affect fish biomass.

Using underwater videos/images to study the survival behaviors of species can inspire new algorithms for solving similar real-world problems, much like the Genghis Khan Shark and Geyser Inspired Algorithms. This multidisciplinary approach will ensure that the study of fish biomass estimation becomes more precise, robust, and applicable to a wider range of environmental conditions, ultimately contributing to the sustainability and efficiency of fish farming practices.

Conflict of interests:

The authors declare that they have no known competing commercial interests or personal relationships that could have appeared to influence the work reported in this research.

Acknowledgments

We want to acknowledge the support provided by Dr. MGR Fisheries College and Research Institute, Ponneri, Chennai, Tamil Nadu. This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 739578, the ADROIT6G project of the SNS-JU under Grant Agreement No. 101095363, and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy. All figures and tables are original creations and not subject to copyright permissions.

Index Of Abbreviations

CLAHE - Contrast Limited Adaptive Histogram Equalization
MAE - Mean Absolute Error
SNR - Signal to Noise Ratio
RMSE - Root Mean Squared Error
MSE - Mean Squared Error
SVM - Support Vector Machine
KNN - k Nearest Neighbour
CNN - Convolutional Neural Network
DNA - Deoxyribonucleic Acid
RCNN - Region Based Convolutional Neural Network
IoU - Intersection over Union

References

Abinaya, N.S., Susan, D. and Sidharthan, R.K., 2022. Deep learning-based segmental analysis of fish for biomass estimation in an occulted environment. Computers and Electronics in Agriculture, 197, p.106985.

Ali, Bulbul & ., Anushka & Mishra, Abha. (2022). Effects of dissolved oxygen concentration on freshwater fish: A review.

Abualigah, L., Al-Ajlouni, Y.Y., Daoud, M.S., Altalhi, M. and Migdady, H., 2024. Fake news detection using recurrent neural network based on bidirectional LSTM and GloVe. Social Network Analysis and Mining, 14(1), pp.1-16.

Agushaka, J.O., Ezugwu, A.E. and Abualigah, L., 2022. Dwarf mongoose optimization algorithm. Computer methods in applied mechanics and engineering, 391, p.114570.

Ancuti, C.O., Ancuti, C., De Vleeschouwer, C. and Bekaert, P., 2017. Color balance and fusion for underwater image enhancement. IEEE Transactions on image processing, 27(1), pp.379-393.

Baidai, Y., Dagorn, L., Amande, M.J., Gaertner, D. and Capello, M., 2020. Machine learning for characterizing tropical tuna aggregations under Drifting Fish Aggregating Devices (DFADs) from commercial echosounder buoys data. Fisheries Research, 229, p.105613.

Barbedo, J.G.A., 2022. A review on the use of computer vision and artificial intelligence for fish recognition, monitoring, and management. Fish, 7(6), p.335.

Benzer, S., Garabaghi, F.H., Benzer, R. and Mehr, H.D., 2022. Investigation of some machine learning algorithms in fish age classification. Fisheries Research, 245, p.106151.

Chen, K., Mahfoud, R.J., Sun, Y., Nan, D., Wang, K., Haes Alhelou, H. and Siano, P., 2020. Defect texts mining of secondary device in smart substation with GloVe and attention-based bidirectional LSTM. Energies, 13(17), p.4522.

Cochrane, K.L. ed., 2002. A fishery manager's guidebook: management measures and their application (No. 424). Food & Agriculture Org.

Coro, G., & Walsh, M. B. (2021, July 1). An intelligent and cost-effective remote underwater video device for fish size monitoring. Ecological Informatics.

Cristianini, N. and Ricci, E., 2008. Support vector machines. In Encyclopedia of Algorithms (pp. 928-932). Springer-Verlag.

Ezugwu, A.E., Agushaka, J.O., Abualigah, L., Mirjalili, S. and Gandomi, A.H., 2022. Prairie dog optimization algorithm. Neural Computing and Applications, 34(22), pp.20017-20065.

Gaude, G.S. and Borkar, S., (2019, May). Fish detection and tracking for turbid underwater video. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 326–331). IEEE.

Hao, Y., Yin, H. and Li, D., (2022). A novel method of fish tail fin removal for mass estimation using computer vision. Computers and Electronics in Agriculture, 193, p.106601.

Hu, G., Guo, Y., Wei, G., & Abualigah, L. (2023, October 1). Genghis Khan shark optimizer: A novel nature-inspired algorithm for engineering optimization. Advanced Engineering Informatics.

Hu, G., Gong, C., Li, X., & Xu, Z. (2024, May 1). CGKOA: An enhanced Kepler optimization algorithm for multi-domain optimization problems. Computer Methods in Applied Mechanics and Engineering.

Hu, J., Zhou, C., Zhao, D., Zhang, L., Yang, G. and Chen, W., (2020). A rapid, low-cost deep learning system to classify squid species and evaluate freshness based on digital images. Fisheries Research, 221, p.105376.

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, 57, p.101088.

Junior, A.D.S.O., Sant'Ana, D.A., Pache, M.C.B., Garcia, V., de Moares Weber, V.A., Astolfi, G., de Lima Weber, F., Menezes, G.V., Menezes, G.K., Albuquerque, P.L.F. and Costa, C.S., 2021. Fingerlings mass estimation: A comparison between deep and shallow learning algorithms. Smart Agricultural Technology, 1, p.100020.

Kumar, V., (2020). Growth and export performance of fish and fish products from India. Indian Journal of Agricultural Marketing, 34(2), pp.15–38.

Li, D., Hao, Y. and Duan, Y., 2020. Nonintrusive methods for biomass estimation in aquaculture with emphasis on fish: a review. Reviews in Aquaculture, 12(3), pp.1390-1411.

Liu, H., Liu, T., Gu, Y., Li, P., Zhai, F., Huang, H. and He, S., (2021). A high-density fish school segmentation framework for biomass statistics in a deep-sea cage. Ecological Informatics, 64, p.101367.

Lines, J.A., Tillett, R.D., Ross, L.G., Chan, D., Hockaday, S. and McFarlane, N.J.B., (2001). An automatic image-based system for estimating the mass of free-swimming fish. Computers and Electronics in Agriculture, 31(2), pp.151–168.

Matveev, V.F. and Steven, A.D.L., 2013. The effects of salinity, turbidity, and flow on fish biomass were estimated acoustically in two tidal rivers. Marine and Freshwater Research, 65(3), pp.267-274.

Pache, M.C.B., Sant'Ana, D.A., Rozales, J.V.A., de Moraes Weber, V.A., Junior, A.D.S.O., Garcia, V., Pistori, H. and Naka, M.H., 2022. Prediction of fingerling biomass with deep learning. Ecological Informatics, 71, p.101785.

Pache, M.C.B., Sant'Ana, D.A., Rezende, F.P.C., de Andrade Porto, J.V., Rozales, J.V.A., de Moraes Weber, V.A., Junior, A.D.S.O., Garcia, V., Naka, M.H. and Pistori, H., 2022. Non-intrusively estimating the live body biomass of Pintado Real® fingerlings: A feature selection approach. Ecological Informatics, 68, p.101509.

Palmer, M., Álvarez-Ellacuría, A., Moltó, V. and Catalán, I.A., 2022. Automatic, operational, high-resolution monitoring of fish length and catch numbers from landings using deep learning. Fisheries Research, 246, p.106166.

Palomares, M.L.D., Froese, R., Derrick, B., Meeuwig, J.J., Nöel, S.L., Tsui, G., Woroniak, J., Zeller, D. and Pauly, D., 2020. Fishery biomass trends of exploited fish populations in marine ecoregions, climatic zones, and ocean basins. Estuarine, Coastal and Shelf Science, 243, p.106896.

Politikos, D.V., Petasis, G., Chatzispyrou, A., Mytilineou, C. and Anastasopoulou, A., (2021). Automating fish age estimation combining otolith images and deep learning: The role of multitask learning. Fisheries Research, 242, p.106033.

Precioso, D., Navarro-García, M., Gavira-O'Neill, K., Torres-Barrán, A., Gordo, D., Gallego, V. and Gómez-Ullate, D., 2022. TUN-AI: Tuna biomass estimation with Machine Learning models trained on oceanography and echosounder FAD data. Fisheries Research, 250, p.106263.

Rao, C.R., and Vinod, H.D., 2019. Conceptual econometrics using R. Elsevier.

Shi, C., Wang, Q., He, X., Zhang, X., and Li, D., (2020). An automatic method of fish length estimation using an underwater stereo system based on LabVIEW. Computers and electronics in agriculture, 173, p.105419.

Sun, M., Hassan, S.G. and Li, D., 2016. Models for estimating feed intake in aquaculture: A review. Computers and Electronics in Agriculture, 127, pp.425-438.

Tengtrairat, N., Woo, W.L., Parathai, P., Rinchumphu, D. and Chaichana, C., 2022. Non-intrusive fish weight estimation in turbid water using deep learning and regression models. Sensors, 22(14), p.5161.

Tolentino, L.K.S., De Pedro, C.P., Icamina, J.D., Navarro, J.B.E., Salvacion, L.J.D., Sobrevilla, G.C.D., Villanueva, A.A., Amado, T.M., Padilla, M.V.C. and Madrigal, G.A.M., 2020. Weight prediction system for nile tilapia using image processing and predictive analysis. International Journal of Advanced Computer Science and Applications, 11(8).

Wageeh, Y., Mohamed, H.E.D., Fadl, A., Anas, O., ElMasry, N., Nabil, A. and Atia, A., 2021. YOLO fish detection with Euclidean tracking in fish farms. Journal of Ambient Intelligence and Humanized Computing, 12, pp.5-12.

Wu, X., Fu, B., Wang, S., Liu, Y., Yao, Y., Li, Y., Xu, Z., & Liu, J. (2023, November 1). Three main dimensions reflected by national SDG performance. The Innovation, Volume 4, Issue 6, 2023, 100507, ISSN 2666-6758

Xuan, K., Deng, L., Xiao, Y., Wang, P. and Li, J., 2023. SO-YOLOv5: Small object recognition algorithm for sea cucumber in complex seabed environment. Fisheries Research, 264, p.106710.

Yassir, A., Andaloussi, S.J., Ouchetto, O., Mamza, K. and Serghini, M., 2023. Acoustic fish species identification using deep learning and machine learning algorithms: A systematic review. Fisheries Research, 266, p.106790.

Beveridge, M.C.M., Phillips, M.J. and Macintosh, D.J. (1997), Aquaculture and the environment: the supply of and demand for environmental goods and services by Asian aquaculture and the implications for sustainability. Aquaculture Research, 28: 797-807 https://doi.org/10.1046/j.1365-2109.1997.00944.x