

# Precision Aquaculture: An Integrated Computer Vision and IoT Approach for Optimized Tilapia Feeding

Rania Hossam<sup>2</sup>, Ahmed Heakl<sup>1</sup> <sup>a</sup> and Walid Gomaa<sup>1,3</sup>

<sup>1</sup>Department of Computer Science and Engineering, Egypt-Japan University of Science and Technology, Alexandria, Egypt

<sup>2</sup>Computer Science Faculty, Mansoura University, Mansoura, Egypt

<sup>3</sup>Faculty of Engineering, Alexandria University, Alexandria, Egypt  
raniaelbadry@std.mans.edu.eg, {ahmed.heakl, walid.gomaa}@ejust.edu.eg

**Keywords:** Computer Vision, Internet of Things (IoT), Aquaculture, Keypoint detection, Object Detection.

**Abstract:** Traditional fish farming practices often lead to inefficient feeding, resulting in environmental issues and reduced productivity. We developed an innovative system combining computer vision and IoT technologies for precise Tilapia feeding. Our solution uses real-time IoT sensors to monitor water quality parameters and computer vision algorithms to analyze fish size and count, determining optimal feed amounts. A mobile app enables remote monitoring and control. We utilized YOLOv8 for keypoint detection to measure Tilapia weight from length, achieving 94% precision on 3,500 annotated images. Pixel-based measurements were converted to centimeters using depth estimation for accurate feeding calculations. Our method, with data collection mirroring inference conditions, significantly improved results. Preliminary estimates suggest this approach could increase production up to 58 times compared to traditional farms. Our models, code, and dataset are open-source: Models - [huggingface.co/Raniahossam33/fish-feeding](https://huggingface.co/Raniahossam33/fish-feeding), Datasets - [huggingface.co/datasets/Raniahossam33/fish\\_feeding](https://huggingface.co/datasets/Raniahossam33/fish_feeding), Code - <https://github.com/ahmedheakl/fish-counting>.


## 1 INTRODUCTION

Optimizing the fish feeding process is critical, as it accounts for up to 40% of total production costs (Atoum et al., 2014; Arditya et al., 2021; Oostlander et al., 2020). Effective nutrient control enhances profitability in aquaculture by preventing waste and maintaining high fish quality. Nutrient wastage not only escalates costs but also contributes to water pollution, adversely affecting fish survival and fertility rates. Therefore, precise nutrient management is essential for both economic efficiency and sustainable aquaculture development, ensuring optimal water quality and operational success.

Recent research has proposed various techniques for controlling the amount of nutrients given to fish. Some researchers have utilized Convolutional Neural Networks (CNNs) for predicting morphological characteristics such as overall length and body size by detecting keypoints on the fish body (Su and Khoshgoftaar, 2009; Tseng et al., 2020). For instance, (Tseng et al., 2020) proposed a CNN classifier to detect only two keypoints, the fish head, and tail fork

regions, to measure the fish body length. Alternatively, (Su and Khoshgoftaar, 2009) used a combination of a faster R-CNN (Ren et al., 2015) for initial fish detection and a stacked hourglass (Newell et al., 2016) for keypoint detection, resulting in a complex and computationally expensive method. Another study (Li et al., 2021) proposed a CNN for marine animal segmentation, which performed well but involved 207.5 million trainable parameters, making it unsuitable for resource-constrained environments like embedded systems or mobile devices.

Automatic control of fish feeding in real environments remains challenging due to variable data appearance and weather conditions, which can affect the accuracy of detection and tracking results (Soetedjo and Somawirata, 2019; Vaquero et al., 2021; Babae et al., 2019). Object tracking, an active research area in computer vision applications, faces increased complexity in multiple-object tracking due to the need for accurate association of objects across frames (Vaquero et al., 2021; Tang et al., 2017; Zhang et al., 2020b). Recent advancements like the SiamRPN tracker (Zhu et al., 2018; Li et al., 2019) and multi-aspect-ratio anchors have significantly improved the performance of Siamese-network-based trackers by

<sup>a</sup>  <https://orcid.org/0009-0009-8712-1457>

addressing the bounding box estimation problem.

To address these challenges, our approach involves the following contributions:

- We develop a method to estimate the weight of Tilapia fish using a length-weight relationship
- We curate an open-source dataset of Tilapia fish images, annotated with keypoints such as the mouth, peduncle, belly, and back.
- We train a YOLOv8 model on this dataset, achieving high precision and recall in keypoint detection and fish counting.
- We design an end-to-end system, powered by two cameras installed in the fish tank, to monitor feeding amounts, pH levels, and dissolved oxygen. The collected data is relayed to a mobile application for easy access and real-time monitoring.

This approach provides a holistic solution for efficient and effective aquaculture management.

The remainder of the paper is organized as follows. Section 2 reviews existing fish mass estimation techniques and feeding methods. Section 3 details our approach, including data collection, model training, and fish weight estimation. Section 4 describes our IoT system architecture for real-time monitoring and control. Results and comparative analysis are presented in Section 5, followed by a discussion of limitations and future work in Section 6. Finally, Section 7 summarizes our findings and their implications for aquaculture productivity.

## 2 RELATED WORK

This section provides an overview of existing research relevant to our study on precise fish feeding in aquaculture. We focus on two key areas: fish mass and length estimation techniques, and automated fish feeding methods. By examining current approaches, we aim to contextualize our work within the field and highlight the advancements offered by our proposed system.

### 2.1 Fish Mass and Length Estimation Techniques

The authors in (Zhang et al., 2020a) developed a fish mass estimation approach by constructing an experimental data collection platform to capture fish images. They used the GrabCut algorithm (Rother et al., 2004) for image segmentation, followed by image enhancement and binarization to extract fish body contours. Shape features were extracted and redundant

features were removed using Principal Component Analysis (PCA) (Maćkiewicz and Ratajczak, 1993), with feature values calculated through a CF-based (Collaborative Filtering) method (Su and Khoshgof-taar, 2009). A BPNN algorithm was then employed to construct the fish mass estimation model. In contrast, our study uses the YOLOv8 model for keypoint detection to identify critical points such as the fish's head and tail. Our dataset is collected using dual-synchronous orthogonal network cameras, with frames analyzed by our backend server. Instead of traditional feature extraction and PCA, we integrate depth estimation using the GLPN (Kim et al., 2022) model to create depth maps, enhancing length measurement accuracy by converting pixel coordinates to real-world dimensions.

The authors in (Jisr et al., 2018; Mathiassen et al., 2011; Islamadina et al., 2018) have used computer vision like saliency map, edge detection and thresholding and traditional image processing techniques like noise reduction and contrast enhancement to segment (Jisr et al., 2018; Islamadina et al., 2018), or make a 3D model (Mathiassen et al., 2011) of the fish body; then use classical machine learning methods, e.g. regression (Sanchez-Torres et al., 2018; Mathiassen et al., 2011) for weight and length extraction. Although these studies have achieved significant results, they have involved complex image processing and feature engineering processes to suit their experimental conditions.

The authors in (Saleh et al., 2023) applied a novel end-to-end keypoint estimation model called MFLD-net. It builds upon CNNs (Sandler et al., 2019), vision transformers (Dosovitskiy et al., 2020), and multi-layer perceptrons (MLP-Mixer) (Tolstikhin et al., 2021). Additionally, it leverages patch embedding (Dosovitskiy et al., 2020), and spatial/channel locations mixing (Tolstikhin et al., 2021). It differs significantly from our approach as the images were taken outside the pool environment. Additionally, their method involved annotating more than four keypoints on each fish, which may increase the complexity of the annotation process. Furthermore, their study used fish of a fixed size, which limits the model's ability to generalize to different fish sizes.

### 2.2 Fish Feeding Techniques

The authors in (Riyandani et al., 2023) focus on developing an automatic feeder employing the YOLOv5x detection model (Vasanthi and Mohan, 2024) for fish feed detection. Their model achieved notable metrics, including an accuracy of 82% and mAP of 81.9%. The automatic feeder dispenses a

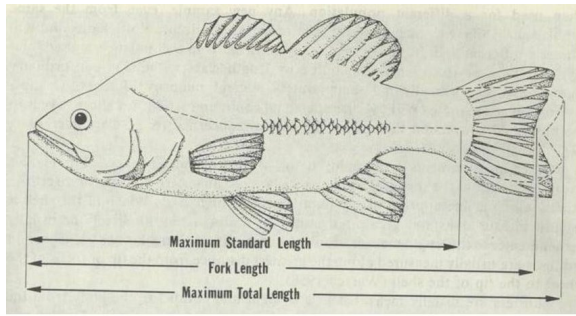


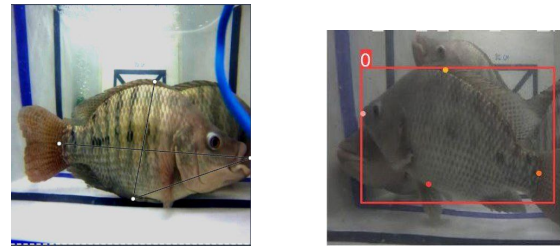
Figure 1: Description of fish lengths. We opt to measure the maximum standard length to calculate fish weight (Froese et al., 2014).

fixed amount of 30 grams of fish feed every five rotations of the stepper motor, with observed variations in fish feed consumption patterns throughout the day. In contrast, our research advances this field by utilizing the more advanced YOLOv8 model for keypoint detection, which promises improved performance. This allows us to adjust the feeding amount dynamically based on real-time measurements of each fish’s size, optimizing feeding practices and preventing overfeeding. Our approach integrates depth estimation to convert 2D image measurements into real-world dimensions, enhancing the accuracy of fish length and weight estimation, and consequently, the daily feeding allowance. This provides a more precise, scalable, and tailored feeding mechanism than the fixed feeding amount used in their study.

The authors in (Tengtrairat et al., 2022) employ a Mask R-CNN (He et al., 2017) with transfer learning to detect Tilapia fish in images. The detection model identifies the fish and extracts dimensions such as length and width. The subsequent weight estimation relies on regression learning models utilizing three key features: fish length, width, and depth. The researchers investigated a regression method for weight estimation like support vector regression (Awad et al., 2015). However, the methodology has several cons. Despite its accuracy, the use of Mask R-CNN is computationally intensive and requires significant processing power. The multi-step process involving depth estimation and feature extraction increases complexity and can be error-prone. Additionally, the regression models, while effective, require precise input features and may not generalize well across varying conditions.

### 3 METHODS

This section details our approach to developing a precise fish-feeding system. We describe the process of



(a) Keypoint annotation. (b) Detection example.

Figure 2: Examples for annotation for fish counting.

estimating Tilapia fish weight, our data collection and annotation methods, the implementation of YOLOv8 for keypoint detection and fish counting, and our technique for calculating fish length and feed amounts. These methods form the foundation of our integrated computer vision and IoT-based solution for optimizing aquaculture management.

#### 3.1 Tilapia Fish Weight Estimation

To determine the appropriate amount of Tilapia fish feed, (Jisir et al., 2018; Lupatsch, 2022) have shown that the amount of feed required for fish can be estimated based on their weight. (M Osman et al., 2020) provided an equation 1 to estimate the fish weight from its length, where the length is defined as the distance from the mouth to the peduncle (Jerry and Cairns, 1998) as shown in figure 1 (Maximum Standard Length):

$$W = aL^b \tag{1}$$

where  $W$  is the fish’s weight in grams (g),  $L$  is the length (cm), and  $a$  and  $b$  are species-specific coefficients ( $a = 0.014$  and  $b = 3.02$  for Tilapia) (M Osman et al., 2020). This method simplifies data collection, bypassing direct weight measurements, and allows weight distribution analysis and other parameters within the fish population.

#### 3.2 Keypoint Annotation and Data Collection

To determine the appropriate feed amount for Tilapia fish, we first needed to specify keypoints to measure their length accurately, defined as the distance from the mouth to the peduncle (Author and Author, Year). For this purpose, we collected 3,500 images of Tilapia fish in a small bowl of three fish. These images were manually annotated using Roboflow (Roboflow, 2022), a widely used tool for creating and managing annotated datasets. Although we only needed the mouth and peduncle keypoints, we annotated four

keypoints on each fish—mouth, peduncle, belly, and back—to aid future research using girth to determine weight (see figure 2a). Following the annotation process, we trained YOLOv8 model (Reis et al., 2023) on the respective dataset to predict the keypoints accurately.

### 3.3 Calculating Fish Length

To estimate fish length, we calculate the Euclidean distance between the head and tail keypoints in pixel units. This measurement is converted to centimeters by integrating depth estimation and focal length, considering the camera distance. Initially, each fish image is resized to a standard ( $416 \times 416$ ) for consistency.

Table 1: Daily feeding allowances as a percentage of fish weight (Riche and Garling, 2003).

Fish Weight Range (g)	Daily Feeding Range (%)
0–1	10 to 30
1–5	6 to 10
5–20	4 to 6
20–100	3 to 4
Larger than 100	1.5 to 3

The GLPN (Global-Local Path Networks) model (Kim et al., 2022) is employed for depth estimation, predicting a depth value for each pixel and creating a depth map essential for spatial information. The YOLOv8 model detects keypoints on the fish (head and tail). These coordinates are then adjusted by their respective depth values to approximate real-world distances. In this depth-adjusted coordinate space, the Euclidean distance between the head and tail keypoints represents the fish's length in pixel units. For the conversion of pixel coordinates to real-world coordinates, given a point in the image with coordinates  $(x_p, y_p)$  and depth  $d$ , the real-world coordinates  $(X, Y, Z)$  can be computed as follows: Let  $f$  be the focal length of the camera (in pixels). Then, we can calculate the 3D coordinates  $(X, Y, Z)$  from the 2D image coordinates  $(x_p, y_p)$  and depth  $d$  as follows:

$$X = \frac{x_p \cdot d}{f}, \quad Y = \frac{y_p \cdot d}{f}, \quad Z = d \quad (2)$$

We can retrieve the length by getting the Euclidean distance between  $(X, Y, Z)_{head}$  and  $(X, Y, Z)_{tail}$  which we call *distance*. Finally, the fish length is calculated using the formula:

$$\text{fish length} = \frac{f}{\text{distance}} \quad (3)$$

### 3.4 Calculating Fish Count

After estimating the fish feed amount based on weight, our next goal is to determine the total feed required for the fish in the bowl. To achieve this, we trained another YOLOv8 model (Reis et al., 2023) on our dataset to count the fish accurately as shown in figure 2b.

### 3.5 Feed Estimation

Once the optimal fish feeding allowances are determined from table 1, we estimate the final feed requirements. By leveraging the robustness of the fish counting models in table 2, the final feed estimation is calculated by multiplying the number of fish, as determined by the fish counting model, with the average feeding amount.

## 4 IoT SYSTEM

Our IoT system integrates a diverse set of sensors including pH, dissolved oxygen (DO), and temperature sensors, along with two cameras, an STM32F103C8 MicroController Unit (MCU) (STMicroelectronics, 2024), and dual pumps—one for feeding fish and another for pH control, as depicted in figure 3. These sensors are connected to the MCU and continuously collect crucial data from the aquatic environment. A prototype is shown in figure 4.

The sensor readings are initially processed by the MCU. Once processed, the MCU transmits the data to gateways within the system architecture. From the gateways, the data is then forwarded via the MQTT communication protocol (Router, 2024) to our backend server. The backend server then acts as the central hub where the data is stored and processed.

The backend server interacts with a dedicated mobile application, shown in figure 6, serving as the user interface. Through this application, users can view real-time graphs, detailed analytics, statistical summaries, and logs reflecting the system's operations and environmental conditions. Simultaneously with sensor data collection, our dual-synchronous orthogonal network cameras actively capture frames from the pool. These frames undergo processing via the MCU, followed by transmission to the gateway, and onward to the MQTT broker before reaching the backend server. At the backend, AI models analyze these frames from two cameras to extract keypoints and fish counts. The results from each camera data and averaged for more accurate predictions.

The AI model's predictions are then relayed back

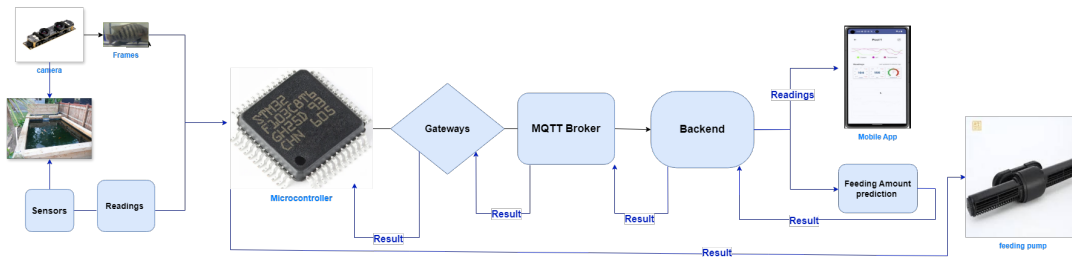


Figure 3: Our IoT system architecture and flow for automated aquarium monitoring and feeding.

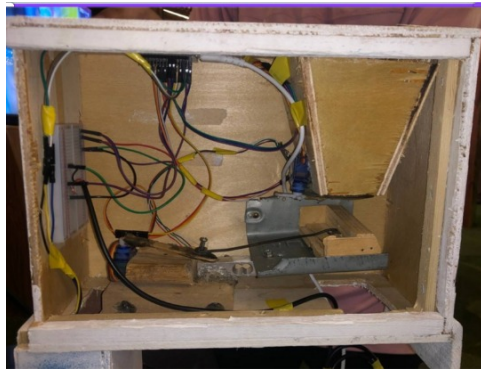


Figure 4: Interior view of the prototype fish feeding system, showing electrical wiring, sensors, and mechanical components housed within a wooden enclosure.

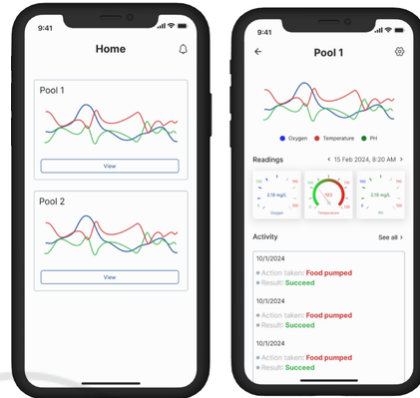


Figure 6: Mobile app readings.

to the backend server, which communicates them via the MQTT broker to the MCU. Based on these insights, the MCU precisely transmits the feeding amount to the feeding pump mechanism. The feeding mechanism operates through a vertical inventory above the pool, regulated by gates, and monitored by a load cell sensor for precise food dispensation. This setup includes a 10KG load cell, an HX-711 amplifier, and two servo motors for meticulous gate control as shown in figure 5.

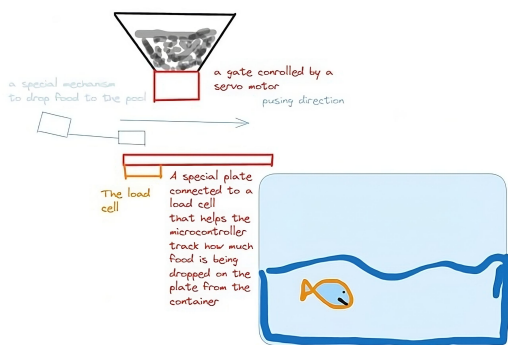


Figure 5: Feeding pump mechanism.

## 5 RESULTS & DISCUSSION

This section presents the outcomes of our experiments using the YOLOv8 model for keypoint detection and fish counting in Tilapia aquaculture. Our evaluation demonstrates that YOLOv8 outperforms existing approaches in both accuracy and speed, making it exceptionally well-suited for deployment on edge devices in resource-constrained environments.

### 5.1 YOLOv8 Performance on Tilapia Dataset

Our experiments demonstrate the superior performance of our YOLOv8 model in both keypoint detection and fish counting tasks. Table 2 summarizes the evaluation metrics for the YOLOv8 model trained on our custom Tilapia fish dataset.

Table 2: Evaluation metrics for keypoints detection and fish-counting models trained on Tilapia fish dataset.

Method	Precision	Recall	AP@50	AP@75
YOLOv8 Keypoints	94.50	89.71	99.68	94.16
YOLOv8 Counting	96.21	86.82	98.88	92.47

For the fish counting task, we employed the same YOLOv8 architecture, which yielded an even higher precision of 96.21%. This precision metric, derived from the model’s training report, represents the ra-

Table 3: Comparison of various scores across different models on the Tilapia dataset on fish Keypoints detection.

Method	Precision	Recall	AP@50	AP@75	AP
Faster R-CNN (Ren et al., 2015)	91.72	85.99	98.50	90.19	67.04
Mask R-CNN (He et al., 2017)	92.61	87.34	99.11	92.12	<b>75.68</b>
RetinaNet (Lin et al., 2017)	90.79	84.26	98.17	83.56	60.53
<b>YOLOv8 (Ours)</b>	<b>94.96</b>	<b>89.06</b>	<b>99.68</b>	<b>94.16</b>	68.04

ratio of correctly detected fish to the total number of detections made by the model. Our fish counting method utilizes frames captured simultaneously from two cameras positioned at different angles in the fish farm. The system processes these paired frames to provide a more comprehensive view of the fish population, helping to reduce occlusions and improve counting accuracy.

To validate these results and assess real-world performance, we conducted additional tests on a separate dataset of 100 frame pairs (200 images total) from actual fish farm conditions. In these tests, the model achieved a counting accuracy of 94.5%, with an average absolute error of 0.7 fish per frame pair. This close alignment between training metrics and real-world performance underscores the model's reliability in practical applications.

The high precision and real-world accuracy of our fish counting model are crucial for accurate population estimation, which directly impacts feed calculation. By combining this accurate count with the average feeding amount determined from our feeding allowance table, we can achieve precise feed estimation, minimizing overfeeding and reducing both feed waste and potential water pollution.

## 5.2 Comparative Analysis with Existing Models

To contextualize our results, we compared YOLOv8's performance with other state-of-the-art deep learning models, as reported in (Tengtrairat et al., 2022). Table 3 presents the comparison of various metrics across different models (Faster R-CNN (Ren et al., 2015), Mask R-CNN (He et al., 2017), RetinaNet (Lin et al., 2017), and our YOLOv8 models (Vasanthi and Mohan, 2024)) on the tilapia dataset.

As evident from table 3, our YOLOv8 model outperforms other models in most metrics, particularly in AP@50 and AP@75. The superior performance in AP@75 is especially noteworthy, as it indicates YOLOv8's ability to maintain high accuracy even with stricter overlap requirements. This is crucial for precise keypoint detection in densely populated fish farms.

While Mask R-CNN shows a higher overall AP

score, which averages performance across all IoU thresholds, YOLOv8 demonstrates more consistent performance at the critical AP@50 and AP@75 levels. This suggests that YOLOv8 may be more reliable for practical applications where moderate to high precision is required.

## 5.3 YOLOv8 & Edge Computing

YOLOv8's architecture is optimized for edge computing, making it ideal for real-time aquaculture monitoring. Its lightweight design allows efficient processing on limited-resource devices, with the nano version achieving sub-200ms inference times on our MCU. This efficiency reduces energy consumption and operational costs. YOLOv8's scalability ensures consistent performance across various hardware configurations, from IoT devices to edge servers. Local data processing minimizes latency and enables rapid decision-making without constant server communication. These features address on-site aquaculture management challenges, potentially revolutionizing Tilapia monitoring. YOLOv8's reliable performance under hardware constraints makes it a superior choice for transforming aquaculture practices.

## 5.4 Implications for Aquaculture Productivity

The high accuracy of our YOLOv8-based system translates to significant potential improvements in aquaculture productivity. Based on preliminary assessments and comparisons with traditional methods, we estimate that our approach can contribute to a 58-fold increase in production compared to conventional fish farms, inspired by (SEAFDEC/AQD, 2022). This dramatic improvement is attributed to:

1. More accurate fish counting, enabling optimal stocking densities.
2. Precise monitoring of fish growth and health through keypoint detection.
3. Reduced water pollution and fish mortality due to timely interventions.

It is important to note that these productivity gains are theoretical maximums based on optimal condi-

tions and full implementation of our system. Real-world results may vary depending on specific farm conditions and management practices.

## 6 LIMITATIONS AND FUTURE WORK

The limitations of this study include the use of datasets from a single fish size in a controlled environment. Future work should include a diverse range of fish sizes and environments to improve model generalizability, especially for smaller fish where keypoint detection is more challenging. Additionally, the current system does not account for varying environmental factors such as water quality, which can influence fish growth and feeding behavior. Integrating environmental monitoring could further optimize feeding practices. While the YOLOv8 model performed well on the Tilapia dataset, its applicability to other fish species remains untested. Lastly, expanding training datasets to include multiple species could enhance its utility across different aquaculture contexts.

## 7 CONCLUSION

This paper used computer vision and IoT technologies to present a novel system for precise Tilapia fish feeding. The system utilizes real-time water quality monitoring and vision-based fish weight estimation to determine optimal feeding amounts. Our models demonstrated superior performance with precision of **94%** for keypoint detection, and **96%** for fish counting, respectively, outperforming Faster R-CNN, Mask R-CNN, and RetinaNet in key metrics. This study provides a precise, scalable solution for sustainable and efficient aquaculture, with recommendations for further real-world testing and refinement. Lastly, this approach has the potential to significantly enhance fish farm productivity (up to 58x) while mitigating environmental concerns by minimizing pollution and fish mortality.

## REFERENCES

Arditya, I., Setyastuti, T. A., Islamudin, F., and Dinata, I. (2021). Design of automatic feeder for shrimp farming based on internet of things technology. *International Journal of Mechanical Engineering Technologies and Applications*, 2(2):145–151.

Atoum, Y., Srivastava, S., and Liu, X. (2014). Automatic

feeding control for dense aquaculture fish tanks. *IEEE Signal Processing Letters*, 22(8):1089–1093.

Author, F. and Author, L. (Year). Standard methods for measuring fish length in ichthyological studies. *Journal of Fish Biology*, XX(X):XX–XX.

Awad, M., Khanna, R., Awad, M., and Khanna, R. (2015). Support vector regression. *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*, pages 67–80.

Babae, M., Li, Z., and Rigoll, G. (2019). A dual cnn-rnn for multiple people tracking. *Neurocomputing*, 368:69–83.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Froese, R., Thorson, J. T., and Reyes Jr, R. B. (2014). Length-weight relationships for 40 species of coral reef fishes from the central philippines. *Journal of Applied Ichthyology*, 30(1):219–220.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

Islamadina, R., Pramita, N., Arnia, F., and Munadi, K. (2018). Estimating fish weight based on visual captured. In *2018 International Conference on Information and Communications Technology (ICOIACT)*, pages 366–372. IEEE.

Jerry, D. and Cairns, S. (1998). Morphological variation in the catadromous australian bass, from seven geographically distinct riverine drainages. *Journal of Fish Biology*, 52(4):829–843.

Jisr, N., Younes, G., Sukhn, C., and El-Dakdouki, M. H. (2018). Length-weight relationships and relative condition factor of fish inhabiting the marine area of the eastern mediterranean city, tripoli-lebanon. *The Egyptian Journal of Aquatic Research*, 44(4):299–305.

Kim, D., Ka, W., Ahn, P., Joo, D., Chun, S., and Kim, J. (2022). Global-local path networks for monocular depth estimation with vertical cutdepth. *arXiv preprint arXiv:2201.07436*.

Li, B., Wu, W., Wang, Q., Zhang, F., Xing, J., and Yan, J. (2019). Siamrpn++: Evolution of siamese visual tracking with very deep networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4282–4291.

Li, L., Dong, B., Rigall, E., Zhou, T., Dong, J., and Chen, G. (2021). Marine animal segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(4):2303–2314.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.

Lupatsch, I. (2022). Optimized feed management for intensively reared tilapia. Accessed: 2024-06-09.

M Osman, H., A Saber, M., A El Ganainy, A., and M Shaaban, A. (2020). Fisheries biology of the haffara bream rhabdosaragus haffara (family: Sparidae) in suez bay,

- egypt. *Egyptian Journal of Aquatic Biology and Fisheries*, 24(4):361–372.
- Maćkiewicz, A. and Ratajczak, W. (1993). Principal components analysis (pca). *Computers & Geosciences*, 19(3):303–342.
- Mathiasen, J. R., Misimi, E., Toldnes, B., Bondø, M., and Østvik, S. O. (2011). High-speed weight estimation of whole herring (*clupea harengus*) using 3d machine vision. *Journal of food science*, 76(6):E458–E464.
- Newell, A., Yang, K., and Deng, J. (2016). Stacked hourglass networks for human pose estimation. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VIII 14*, pages 483–499. Springer.
- Oostlander, P., van Houcke, J., Wijffels, R. H., and Barbosa, M. J. (2020). Microalgae production cost in aquaculture hatcheries. *Aquaculture*, 525:735310.
- Reis, D., Kupec, J., Hong, J., and Daoudi, A. (2023). Real-time flying object detection with yolov8. *arXiv preprint arXiv:2305.09972*.
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- Riche, M. and Garling, D. (2003). Feeding tilapia in intensive recirculating systems.
- Riyandani, R., Jaya, I., and Rahmat, A. (2023). Computer vision-based fish feed detection and quantification system. *Journal of Applied Geospatial Information*, 7(1):832–839.
- Roboflow (2022). Roboflow. <https://roboflow.com/>.
- Rother, C., Kolmogorov, V., and Blake, A. (2004). ” grab-cut” interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, 23(3):309–314.
- Router, O. (2024). What is mqtt? Accessed: 2024-05-31.
- Saleh, A., Jones, D., Jerry, D., and Azghadi, M. R. (2023). Mfld-net: a lightweight deep learning network for fish morphometry using landmark detection. *Aquatic Ecology*, 57(4):913–931.
- Sanchez-Torres, G., Ceballos-Arroyo, A., and Robles-Serrano, S. (2018). Automatic measurement of fish weight and size by processing underwater hatchery images. *Engineering Letters*, 26(4):461–472.
- Sandler, M., Baccash, J., Zhmoginov, A., and Howard, A. (2019). Non-discriminative data or weak model? on the relative importance of data and model resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0.
- SEAFDEC/AQD (2022). Farmers must understand the costs of tech-driven aquaculture. Accessed: 2024-06-06.
- Soetedjo, A. and Somawirata, I. K. (2019). Improving traffic sign detection by combining msr and lucas kanade tracking. *IMPROVING TRAFFIC SIGN DETECTION BY COMBINING MSER AND LUCAS KANADE TRACKING*, 15(2):1–13.
- STMicroelectronics (2024). Stm32f103c8 microcontroller. Accessed: 2024-05-31.
- Su, X. and Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009(1):421425.
- Tang, S., Andriluka, M., Andres, B., and Schiele, B. (2017). Multiple people tracking by lifted multicut and person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3539–3548.
- 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):5161.
- Tolstikhin, I. O., Houlsby, N., Kolesnikov, A., Beyer, L., Zhai, X., Unterthiner, T., Yung, J., Steiner, A., Keysers, D., Uszkoreit, J., et al. (2021). Mlp-mixer: An all-mlp architecture for vision. *Advances in neural information processing systems*, 34:24261–24272.
- Tseng, C.-H., Hsieh, C.-L., and Kuo, Y.-F. (2020). Automatic measurement of the body length of harvested fish using convolutional neural networks. *Biosystems Engineering*, 189:36–47.
- Vaquero, L., Mucientes, M., and Brea, V. M. (2021). Siammt: Real-time arbitrary multi-object tracking. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 707–714. IEEE.
- Vasanthi, P. and Mohan, L. (2024). Multi-head-self-attention based yolov5x-transformer for multi-scale object detection. *Multimedia Tools and Applications*, 83(12):36491–36517.
- Zhang, L., Wang, J., and Duan, Q. (2020a). Estimation for fish mass using image analysis and neural network. *Computers and Electronics in Agriculture*, 173:105439.
- Zhang, Y., Wang, C., Wang, X., Zeng, W., and Liu, W. (2020b). A simple baseline for multi-object tracking. *arXiv preprint arXiv:2004.01888*, 7(8).
- Zhu, Z., Wang, Q., Li, B., Wu, W., Yan, J., and Hu, W. (2018). Distractor-aware siamese networks for visual object tracking. In *Proceedings of the European conference on computer vision (ECCV)*, pages 101–117.