A Computer Vision Based Behavioral Study and Fish Counting in a Controlled Environment

KRITHIKA M. PAI, K. B. AJITHA SHENOY, AND M. M. MANOHARA PAI, (Senior Member, IEEE)
Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India
Corresponding authors: K. B. Ajitha Shenoy (ajith.shenoy@manipal.edu) and M. M. Manohara Pai (mmm.pai@manipal.edu)

ABSTRACT
Aquaculture is a fast-growing food-production sector that accounts for almost 50% of the world’s fish used for consumption. Aquaculture refers to the cultivation of fish in cages. The fish in the cage must be fed at regular intervals, for which the number of fish helps in estimating the amount of feed to be put in the cage. The behavior of fish in a caged environment reflects their health. In the absence of an ambient atmosphere, fish are stressed, which results in frantic movement. The frantic behavior of fish can be identified using recent advancements in image and video processing. In this study, we have focused on frantic behavior detection, fish detection, and counting. For this, a RAS with Tilapia fish has been setup, and the videos of the fish are captured. The detection and counting have been achieved by using the YOLOv5 model. The model has resulted in a Precision, Recall and F-measure of 81%. The results are compared with the ground truth, which indicates that the model has been successful in counting the fish. The frantic movement of the fish has been detected by developing an optical flow model. The results are encouraging and can be used for frantic behavior detection.

INDEX TERMS
Fish counting, frantic behavior, detection.

I. INTRODUCTION
India is the second-largest producer of aquaculture fish and the third-largest fish-producing country globally. Fisheries and aquaculture are essential for poverty alleviation. Fish accounts for 17% of the animal protein intake [1] globally and remains one of the most traded food commodities worldwide. By 2050 around 125-210 million tons of fish are expected to be required to meet the annual per capita requirement of 15-20 kg [2]. Aquaculture refers to the cultivation of aquatic organisms in a controlled environment. Aquaculture has been growing at a rapid pace and is part of global food security and sustainable development growth. It has gained immense importance in recent times as fish consumption is increasing by 6% annually. The demand has doubled since the beginning of year 2000 and will double again if the current growth rate holds. Aquatic animals, especially fish, are a renewable resource that people can feed on indefinitely if managed sustainably.

Cage aquaculture involves growing fish in their natural environment while being enclosed in a cage that allows the free flow of water. The fish be it in cage, pond or RAS (Recirculatory Aquaculture System) system must be fed at regular intervals. Since 60% of the operating expenses are used for feeding [3], it is important to optimize the use of fish feed to avoid feed wastage and contamination of water inside the cage. The counting of fish is also helpful in determining the mortality rate of the fish. It is challenging to keep track of fish and their behavior manually inside a cage as several behavioral changes can be seen in cultivated fish [4]. It costs a lot of time and money. Monitoring of fish in the aquaculture systems [5] is essential to ensure that the fish are not stressed and have ambient surroundings to grow.

The fish tend to move faster when stressed and this movement is referred to as frantic movement. A frantic movement pattern observed in the absence of the feed can be due to stress. The stress in fish can be detected through their
movement in the water. Physical, chemical, and perceived stressors can evoke a non-specific response in fish [6]. Bad water quality also affects the behavior of the fish. Monitoring or predicting the water quality parameters is very important in behavioral study [7], [8]. Stress can affect the fish’s reproduction, growth, and resistance [9]. Early detection can help in better growth and development of the fish. The advancement in technology gives us hope to conduct such experiments more straightforwardly.

The experiment in this work has been performed on Tilapia fish. These are edible freshwater fish having a high nutritional value. These are commonly consumed fish worldwide and is the most farmed fish species in over 120 countries and territories across the globe [10]. Tilapia fish are fresh water fish found in ponds and shallow waters. These fish are very much affordable and provide a great amount of nutritional value as they are a good source of omega-3 fatty acids [11] that make up good fat. These are also beneficial for cancer prevention and good heart health.

RAS [12], [13], [14] has been used to cultivate these fish. In RAS, the water is recycled and reused after mechanical and biological infiltration. New water is added to the system to make up for the splash-out, evaporation, and any other process that leads to water loss from the system. RAS has several advantages such as better fish quality, close monitoring, risk reduction due to climatic factors, disease, and parasite impacts.

The contributions of this paper are:
- Fish detection and counting
- An algorithm for frantic movement pattern detection

The paper is organized as follows: Section I gives the Introduction to the research work. Section II shows the Related Works. The Experimental Setup, Dataset Preparation and Proposed Methodology are given in Section III, Section IV, Section V respectively. Section VI shows the results obtained from this research. Finally the paper is concluded in Section VII.

II. RELATED WORKS
Shreesha et al. [15] have performed a behavioral analysis of Silago sihama fish in an aquarium setup. An analysis is performed on the various types of anomalous behaviors of the fish, such as swimming at the surface, frantic movement behavior, and no movement, which has been mathematically computed. Wang et al. [16] have used a convolutional neural network, YOLOv2, to classify and detect the fish images. They have used this model to detect the fish and classified them into one of the seven classes, such as shark, dolphinfish, yellowfin tuna etc. Lumauag et al. [17] have proposed a method to count and track the fish, which was carried out using blob analysis and Euclidean filtering. The authors claim that the model exhibits an average counting accuracy of 91% and an average detection accuracy of 94%. Hossain et al. [18] have worked on detecting the underwater moving objects as fish, identifying the species of those fish, and tracking the detected fish to avoid multiple counting. GMM-based model background subtraction method has been used to detect objects and the Kalman filter for object tracking. Pyramid Histogram of Visual Words (PHOW) is used as a feature vector and SVM as a classifier. Kong et al. [19] have done a comparative study of fish detection wherein they have modified the YOLOv5 algorithm to achieve better performance. The authors have made use of 3 models, namely Centernet, YOLOv4, and YOLOv5, for Golden Crucian Carp detection. They have shown that the YOLOv5 model has performed better than all other models.

Mohamed et al. [20] have tracked the movement of the fish by making use of optical flow algorithm and combination of YOLO + optical flow. They have concluded that the results obtained from YOLO + optical flow were much better than the optical flow algorithm only as YOLO + optical flow was able to track 4/4 fish whereas only optical flow could track 1/4 fish only.

Sung et al. et al. [21] have made use of YOLO algorithm and have compared it with the HOG classifier-based algorithm and have proved that the former has better performance than the latter as it can successfully process noisy, dim light, and hazy underwater image.

Naseer and Baro [22], [23] have come up with a refinement technique based on spatial temporal analysis to improve the mAP for detection of Nephrops burrows. The proposed refinement technique is said to work for generic detection models. From this Literature Review it is evident that YOLO models give accurate detection results and hence YOLOv5 has been implemented in this work for detection of fish. To the best of our knowledge there has been no work on frantic behavior detection. So in this work an algorithm has been developed to detect frantic behavior pattern in the video sequences.

III. EXPERIMENTAL SETUP
A tank-based RAS (Recirculatory Aquaculture System) [24] of dimensions 8 feet × 6 feet × 4 feet, as shown in Fig.1 has been set up in the Institute premises for the experimental purpose and consists of Tilapia fish. The details of the experimental setup and dataset are shown in Table 1. It has been constructed using cement. A few portions of the pond are covered with glass for better visibility from the sides. The bottom of the system is covered with sandbed. An aerator is used to mix air with water. The water in the RAS is recirculated through the biofilter. The biofilter allows the water to pass through it. This filter serves as a home for bacteria that break down fish waste to keep the environment safe and non-toxic. The water quality in the pond is closely monitored using water quality sensors as shown in Table 2 so that the fish are provided with a conducive ecosystem.

IV. DATASET PREPARATION
An underwater camera is used for capturing the underwater video. Two videos are captured and converted into frames. The annotations have been done using Roboflow which acts
V. PROPOSED METHODOLOGY

The video data is manually collected from the RAS system using an underwater camera. The videos captured are then annotated and used for training the detection model. Fish counting is then performed on the detected images. For frantic movement detection, the fish are tracked across the frames. The displacement of the fish in the successive frames are calculated and stored which are then averaged to find a global threshold $\tau$. Once the threshold $\tau$ is defined, the frames with values greater than $\tau$ are considered to be frantic. The block diagram for the work is shown in Fig. 2.

A. FISH TRACKING AND DETECTION

YOLO [25], which stands for 'You Only Look Once' is an object detection algorithm used for object detection and uses a Convolutional Neural Network (CNN) [26]. This algorithm has gained a lot of popularity due to its high-speed accuracy and learning capabilities. The algorithm detects object through single forward propagation.

The YOLOv5 network architecture as shown in Fig. 3 consists of three parts namely Backbone, Neck and Head where CSPDarknet [27] makes up the Backbone and is used for feature extraction, PANet [28] makes up the neck and is essential for feature fusion and the head is made up of Yolo which is used for the detecting output results. YOLOv5 model, has been trained for the detection of fish. It has been trained using Roboflow. A total of 508 images has been used. The ratio of train, validation, and test is set to 70%, 15%, and 15%, respectively. The images are resized to $416 \times 416$.

B. OPTICAL FLOW

Optical flow [30] is a motion estimation technique and is usually applied to the images having a small time step between them. The velocity for points within the images is calculated. This work captures the displacement of the fish detected by the bounding boxes. The equations (1)-(4) are used for optical flow.

Algorithm 1 Fish Detection and Counting

- Step 1: Collect the video of fish using underwater camera
- Step 2: Convert the video into frames
- Step 3: Annotate the fish in each of the frame
- Step 4: Split the frame set in train, validation and test set
- Step 5: Train the YOLOv5 model using the train set
- Step 6: Use the trained model to detect and count the fish in the test set

Algorithm 2 Frantic Behaviour Detection

- Step 1: Collect the video of fish using underwater camera
- Step 2: Convert the video into shorter video sequences
- Step 3: Classify the video sequences captured in Step 2 into Normal or Frantic based on observation
- Step 4: Apply the optical flow on the Normal sequences
- Step 5: Obtain the movement of fish in consecutive frames
- Step 6: Find the threshold using the Otsu method for each Normal video sequence
- Step 7: Average the thresholds obtained in Step 6 to obtain a global threshold $\tau$
- Step 8: Classify the video sequences as frantic or not using the global threshold.
**FIGURE 2.** Block diagram for the research work.

**FIGURE 3.** YOLOv5 architecture [29].
FIGURE 4. Training graphs obtained for YOLOv5 model.

### TABLE 3. Evaluation metrics.

| mAP  | Precision | Recall | F-measure |
|------|-----------|--------|-----------|
| 0.86 | 0.81      | 0.81   | 0.81      |

**C. THRESHOLDING**

To classify the fish movement as normal or frantic, a threshold has been found out. Otsu method [32] is used for automatic image thresholding. It maximizes the between class variance. The threshold values for all the video sequences are obtained and averaged which is the global threshold, \( \tau \).

**VI. RESULTS AND DISCUSSION**

#### A. FISH TRACKING AND DETECTION

YOLOv5 model has been implemented for fish detection and counting. Based on the Precision, recall and F-measure, the results obtained for fish detections can be considered as good. The model is able to detect the fish almost accurately.

In order to measure the performance of the model, mean Average Precision (mAP), Precision(5), Recall(6) and F-score(7) have been calculated.

Precision is calculated as the number of true positives divided by the total number of true positives and false positives.

\[
\text{Precision} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}} \tag{5}
\]

Recall is calculated as the number of true positives divided by the total number of true positives and false negatives.

\[
\text{Recall} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}} \tag{6}
\]

F-Measure presents a single score that evaluates both the concerns of precision and recall as shown in

\[
F - \text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}
\]

The results of optical flow for one frame are shown in the Fig. 6. The evaluation metrics are shown in Table 3.

#### B. OPTICAL FLOW

The video sequences are manually classified as normal or frantic based on the displacement of fish. Since no prior research has been conducted on the frantic movement pattern of the fish, it has been done manually to establish a ground-truth. 10 video sequences have been classified as normal and 6 video sequences have been classified as frantic based on observation of the fish movement in the video sequence.

Optical flow is then applied on the normal video sequences. The displacement of the fish in the consecutive frames are noted in a csv file.

The results of optical flow for one frame are shown in the Fig. 7.
C. THRESHOLDING

Otsu method has been applied to all the csv files obtained in the previous step. The obtained thresholds are then averaged to find a global threshold $\tau$ for the frantic movement classification as shown in Table 4. Here the value of $\tau$ is found to be 0.0455.

Optical flow is then applied on the frantic video sequences with the threshold set to $\tau$. The frames having displacement values greater than $\tau$ are classified as frantic frames and the those with displacement values lower than $\tau$ are classified as normal frames. The percentage of frantic frames are the calculated using (8), as shown at the top of page 8.
FIGURE 6. Ground truth and model predictions showing the case where the model fails to detect two separate fish due to occlusion. Fish class is encoded as 0 in the above figure.

FIGURE 7. Optical flow outputs.

The percentages of frantic frames in each of the video sequence are shown in Table 5.

In this particular research work, fish detection and counting and frantic behavior detections have been performed. Decent results were obtained from YOLOv5 model in-terms of accuracy, precision, recall and F-measure for fish detection and counting.

Early detection of frantic movement in fish facilitates fish farmers to take precautionary measures promptly. The percentages of frantic frames in a video sequence gives a
general idea of the pattern of fish movement. In this particular research, the video sequences for frantic behavior detection are very short (10 sec). A larger percentage of frantic frames, if detected in a longer video sequence, then the video sequence itself can be classified as a frantic video sequence. If there are multiple video sequences which are classified as frantic, immediate measures have to be taken in order reduce the fish stress.

VII. CONCLUSION
Aquaculture has gained immense importance in the recent times as it has increased the number of potential jobs as it produces a product for the marketplace as well as creates employment opportunities. It is a source of food as well as an extra source of income. This work has focused on the detections and counting of fish and frantic movement behaviour detections in the aquaculture setup. The early detection of frantic behaviour pattern is essential to ensure that the fish in the aquaculture system are provided with an ambient environment for growth. In this work, the fish detection and counting has been carried out using the YOLOv5 model. The model has managed to detect and count a good percentage of fish in the frame. The detected fish has been tracked using Optical flow and Otsu thresholding method has been used to find the global threshold. The threshold has then been used to find the percentage of frantic frames in the video sequence and thus classify the video sequences as normal or frantic.

ACKNOWLEDGMENT
The authors would like to thank S. Shreesha, a Research Scholar for assisting whenever necessary during the course of this research work and also would like to thank Dr. Radhika M. Pai, a Professor and the Head of the Department of Data Science and Computer Applications for providing experimental setup and laboratory resources.

REFERENCES

[1] C. Béné, M. Barange, R. Subasinghe, P. Pinstrup-Andersen, G. Merino, G.-I. Hemre, and M. Williams, “Feeding 9 billion by 2050—putting fish back on the menu,” Food Secur., vol. 7, pp. 261–274, Mar. 2015.
[2] J. P. Fry, N. A. Mailoux, D. C. Love, M. C. Milli, and L. Cao, “Feed conversion efficiency in aquaculture: Do we measure it correctly?” Environ. Res. Lett., vol. 13, no. 2, Feb. 2018. Art. no. 024017.
[3] B. Arru, R. Furesi, L. Gasco, and F. A. Madau, “Insect meal and fish feeding cost: A case study in Italy,” in Green Metamorphoses: Agriculture, Food, Ecology, Wageningen, The Netherlands: Wageningen Academic, 2020.
[4] K. Barrington, “Behavioral changes and problems in aquarium fish,” Tech. Rep., 2022. [Online]. Available: https://www.ratemyfishtank.com/blog/behavioral-changes-and-problems-in-aquarium-fish
[5] M. Holmer, P. Hansen, I. Karakassis, J. A. Borg, and P. Schembri, Monitoring of Environmental Impacts of Marine Aquaculture. Dordrecht, The Netherlands: Springer, Dec. 2007, pp. 47–85.
[6] B. A. Barton. “Stress in fishes: A diversity of responses with particular reference to changes in circulating corticosteroids,” Integrative Comparative Biol., vol. 42, pp. 517–525, Jul. 2002.
[7] S. Girisha, M. M. M. Pai, U. Verma, R. M. Pai, and S. Shreesha, “Anomaly detection using classification CNN models: A video analytic approach,” in Proc. IEEE Region 10 Conf. (TENCON), Dec. 2021, pp. 923–928.
[8] S. Shreesha, M. M. M. Pai, R. M. Pai, and U. Verma, “LSTM-based prediction of water quality parameters system in backwaters,” in Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONNIEC), Jul. 2021, pp. 1–6.
[9] M. Svoboda. “Stress in fish: A review,” in Research Institute of Fish Culture and Hydrobiology, vol. 37. Vertebratae, 2001, pp. 169–191.
[10] A. K. Claire Sissons. “Is tilapia healthy? Safety to eat, nutrition, and how farmers raise it,” Tech. Rep., 2021. Art. no. 322493. [Online]. Available: https://www.medicalnewstoday.com/articles/322493#safety
[11] K. Young. “Omega-6 (n-6) and omega-3 (n-3) fatty acids in tilapia and human health: A review,” Int. J. Food Sci. Nutrition, vol. 60, pp. 203–211, Jan. 2009.
[12] Recent Trends in Aquaculture Recirculatory Aquaculture System (RAS), National Fisheries Development Board, Animal Husbandry & Dairying, Government of India, Hyderabad, India. [Online]. Available: https://static.vikaspedia.in/mediastorage/document/Ras_Booklet_Eng.pdf
[13] L. A. Hellfrich and G. Libey, Fish Farming in Recirculating Aquaculture Systems (RAS). Blacksburg, VA, USA: Virginia Cooperative Extension, 1991.
[14] E. Schoina, A. I. Doulgeraki, H. Miliou, and G.-J.-E. Nychas, “Dynamics of water and biofilm bacterial community composition in a Mediterranean recirculation aquaculture system,” Aquaculture J., vol. 2, no. 2, pp. 164–179, May 2022.
[15] S. Shreesha, M. M. M. Pai, U. Verma, and R. M. Pai, “Computer vision based fish tracking and behaviour detection system,” in Proc. IEEE Int. Conf. Distrib. Comput., VLSI, Electr. Circuits Robot. (DISCOVER), Oct. 2020, pp. 252–257.
[16] M. Wang, M. Liu, F. Zhang, G. Lei, J. Guo, and L. Wang, “Fast classification and detection of fish images with YOLOv2,” in Proc. OCEANS-MTS/IEEE Kobe Techno-Oceans (OTCOS), May 2018, pp. 1–4.
[17] R. Lumaan and M. Nava, “Fish tracking and counting using image processing,” in Proc. IEEE 10th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ. Manag. (HNICEM), Nov./Dec. 2018, pp. 1–4.
[18] E. Hossain, S. M. S. Alam, A. A. Ali, and M. A. Amin. “Fish activity tracking and species identification in underwater video,” in Proc. 5th Int. Conf. Informat., Electron. Vis. (ICIENV), May 2016, pp. 62–66.
[21] M. Sung, S.-C. Yu, and Y. Girdhar, “Vision based real-time fish detection using convolutional neural network,” in *Proc. OCEANS-Aberdeen*, Jun. 2017, pp. 1–6.

[22] A. Naseer, E. N. Baro, S. D. Khan, and Y. Vila, “A novel detection refinement technique for accurate identification of nephrops norvegicus burrows in underwater imagery,” *Sensors*, vol. 22, no. 12, p. 4441, 2022.

[23] A. Naseer, E. N. Baro, S. D. Khan, and Y. V. Gordillo, “Automatic detection of nephrops norvegicus burrows in underwater images using deep learning,” in *Proc. Global Conf. Wireless Opt. Technol. (GCWOT)*, Oct. 2020, pp. 1–6.

[24] C. A. Espinal and D. Matulić, *Recirculating Aquaculture Technologies*. Cham, Switzerland: Springer, 2019, pp. 35–76.

[25] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” 2015, arXiv:1506.02640.

[26] K. O’Shea and R. Nash, “An introduction to convolutional neural networks,” 2015, arXiv:1511.08458.

[27] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, “YOLOv4: Optimal speed and accuracy of object detection,” 2020, arXiv:2004.10934.

[28] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path aggregation network for instance segmentation,” 2018, arXiv:1803.01534.

[29] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, “A forest fire detection system based on ensemble learning,” *Forests*, vol. 12, no. 2, p. 217, Feb. 2021.

[30] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.

[31] Wikipedia Contributors. (2022). *Optical Flow—Wikipedia, the Free Encyclopedia*. Accessed: Jul. 19, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Optical_flow&oldid=1075679216

[32] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.