DeepFoids: Adaptive Bio-Inspired Fish Simulation with Deep Reinforcement Learning (Appendix)

Yuko Ishiwaka[1]*[†], **Xiao S. Zeng**[2]*, **Shun Ogawa**[1],
Donovan Michael Westwater[2], **Tadayuki Tone**[1], **Masaki Nakada**[2][‡]
[1]SoftBank Corp., Japan, [2] NeuralX Inc., USA

A Biological Background

Fish species prefer different temperatures and will change their swimming patterns or vertical alignment based on the surrounding water temperature [1]. Generally, adult salmonids prefer water temperature within the 8-9 degrees Celsius range [2], whereas yellowtail amberjack and red seabream like higher temperatures [3, 4]. Caged fish also show distinct swimming patterns at daytime and night. They move very little and stay closer to water surface after sunset, but gradually become more active and swim deeper after sunrise as the amount of sunlight increases [5, 6]. Additionally, when fish in the cage reach a certain crowding density threshold, they attempt to maintain their optimal personal space while still being able to move around [7] by forming a circular schooling pattern, consistent with the Boids model [8]. Fish normally react to their neighbors within a distance of 2 to 3 body lengths (BLs) [9, 10] and can learn to avoid cage walls [11, 12]. The reaction appears regulated by visual stimuli [7], as well as the lateral line system, which allows the fish to feel the vibrations of their neighbors [13]. Fish align their orientation primarily to the fish in front of them, and adjust speed as necessary to avoid collision with neighbors [10]. Caged fish like salmonids often settle into two or three vertical layers [14] since they typically group with similarly sized fish [15], where larger fish tend to swim at lower depths and smaller fish stay at higher depths [16, 17]. Moreover, the decision making interval of 95% of fish species is controlled by the Mauthner cells (M-cells) that trigger a rapid escaping movement (C-start) for risk avoidance and command neurons active in decision-making [18, 19]. Miller et al. [20] studied zebrafish social groups and found that action potentials for swimming, which is generated by the neural circuit, lasted for a duration of 50-200 ms. Since salmonids have lower body temperature whereas yellowtail amberjack and red seabream have similar body temperatures as zebrafish, we set 200ms as the the decision-making interval for coho salmon and 100ms as that for yellowtail amberjack and red seabream.

B Environmental Background

Colors underwater are determined by the intrinsic colors of objects and the scattering and absorption of light. The main factors that determine the absorption and scattering are the properties of the water itself and the particulates in the water, usually chlorophyll and sediments [21]. Water absorbs light of different wavelengths at different rates, characterized by attenuation coefficients, which can be changed by the presence of particulates [22, 23]. The exact relationship between transmitted light and the light absorbed is expressed by the Beer Lambert law [21]. Chlorophyll absorbs blue light the most, followed by red, then green. Thus a higher concentration of chlorophyll shifts the coloring in a reddish green direction. Sediments in the water scatter the light, altering the amount of light absorbed

---

*Yuko Ishiwaka and Xiao S. Zeng equally contributed to this work.
[†]yuko.ishiwaka@g.softbank.co.jp
[‡]masaki@neuralx.ai

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

as well as changing the attenuation coefficients. Sediments and suspended solids also create turbidity and occlusion underwater. In our model, sediment concentration is the main driving force behind turbidity.

C Simulation Setting

We created eleven simulation environments to assess the performance of proposed fish simulation framework and generate synthetic dataset for computer vision tasks. The configuration of each environment is available in Table 1. The first three environments reproduce the underwater scenes in fish farms we collected field data from and are used as default scenes to pretrain each of the three species. The other eight environments are used to examine the adaptiveness of the framework and transfer learning is applied. Values in the cage size column are in units of meter. The biological and environmental parameters used in these scenes are summarized in in Table 2.

Table 1: Simulation environment configurations

Species	Fish Number	Body Scale	Cage Size	Cage Shape
Coho salmon	1000	[0.9, 1.1]	Edge of 3, Height of 4.6	Octagon
Yellowtail amberjack	45	[0.9, 1.1]	3×3×3	Cube
Red seabream	10	[0.9, 1.1]	3×3×3	Cube
Red seabream	10	[0.9, 1.1]	5×5×5	Cube
Red seabream	50	[0.9, 1.1]	3×3×3	Cube
Coho salmon	300	0.5	Edge of 3, Height of 4.6	Octagon
Coho salmon	300	1.0	Edge of 3, Height of 4.6	Octagon
Coho salmon	300	1.5	Edge of 3, Height of 4.6	Octagon
Coho salmon	1000	0.5, 1.0 or 1.5	Edge of 1.8, Height of 3	Octagon
Coho salmon	1000	0.5, 1.0 or 1.5	Edge of 3, Height of 4.6	Octagon
All three species	300	[0.9, 1.1]	Edge of 3, Height of 4.6	Octagon

D Effects of Light Intensity and Temperature on Fish Behavior

A fish prefers to stay within a certain area of the tank where its light intensity preference matches the local light intensity computed by the Beer-Lambert law:

$$I = I_0 \times e^{-ad} \tag{1}$$

where I and I_0 denote the light intensity at the local depth and on water surface, a is an attenuation coefficient and d is the depth of fish. I_0 and α are valued based on the field data collected from aquaculture sites as described in Sec. I. Any deviation of I from the light intensity preference range results in a vertical delta velocity ($\Delta\mathbf{v}_{\text{light}}$) for the fish as shown in Eq. (2):

$$\Delta\mathbf{v}_{\text{light}} = \begin{cases} \frac{I_{\text{lpref}} - I}{I_{\text{lpref}} - I_{\text{lsteep}}} [0, 1, 0] & \text{if} \quad I \leq I_{\text{lpref}} \\ \frac{I - I_{\text{hpref}}}{I_{\text{hsteep}} - I_{\text{hpref}}} [0, -1, 0] & \text{if} \quad I \geq I_{\text{hpref}} \end{cases}, \tag{2}$$

where parameters I_{lsteep} and I_{hsteep} control the steepness of the reaction to light and are set in an identical manner to [9]. I_{lpref} and I_{hpref} denote the lower and upper bounds of the light preference interval and are valued based on the biological studies of each simulated species [24, 3, 4, 5]. Similarly, a fish likes to stay in the area where the local temperature satisfies its temperature preference. We calculate the temperature induced vertical delta velocity ($\Delta\mathbf{v}_{\text{temp}}$)

$$\Delta\mathbf{v}_{\text{temp}} = \begin{cases} \frac{T_{\text{lpref}} - T}{T_{\text{lpref}} - T_{\text{lsteep}}} [0, 1, 0] & \text{if} \quad T \leq T_{\text{lpref}} \\ \frac{T - T_{\text{hpref}}}{T_{\text{hsteep}} - T_{\text{hpref}}} [0, -1, 0] & \text{if} \quad T \geq T_{\text{hpref}} \end{cases}, \tag{3}$$

where the under-water ambient temperature (T) is obtained using a two-layer method based on the studies by [25] and [26]. The temperature response steepness parameters (T_{lsteep} and T_{hsteep}) are set to be the same as those in [9], and the bounds of the temperature preference interval (T_{lpref} and T_{hpref}) are determined according to [3], [4] and [2].

2

Table 2: Biological and environmental parameters in fish simulation

Description	Parameters	Unit	Value
Body length	BL	m	[0.34, 0.52]
Speed while schooling (coho salmon)	v_t^a	BL/second	[0.2, 1.9]
Speed while schooling (yellowtail amberjack)	v_t^a	BL/second	[0.2, 2.1]
Speed while schooling (red seabream)	v_t^a	BL/second	[0.2, 2.7]
Initial speed	v_0	BL/second	0.2
Maximum delta speed in cage environment	$\Delta v_{\max}$	BL/second	4.1
Light intensity at water surface (Onmaehama)	I_0	PAR	[0, 519.69]
Light intensity at water surface (Nishiki)	I_0	PAR	[0, 15.13]
Light attenuation coefficient (Onmaehama)	a	N/A	1.5
Light attenuation coefficient (Nishiki)	a	N/A	0.26
Lower bound of preferred light intensity (coho salmon, regular size)	I_{lpref}	PAR	[1.1, 4.5]
Lower bound of preferred light intensity (coho salmon, small size)	I_{lpref}	PAR	[1.1, 4.5]
Lower bound of preferred light intensity (coho salmon, large size)	I_{lpref}	PAR	[0.1, 0.5]
Lower bound of preferred light intensity (yellowtail amberjack)	I_{lpref}	PAR	[0.01, 0.02]
Lower bound of preferred light intensity (red seabream)	I_{lpref}	PAR	5.75
Upper bound of preferred light intensity (coho salmon, regular size)	I_{hpref}	PAR	[3.5, 7]
Upper bound of preferred light intensity (coho salmon, small size)	I_{hpref}	PAR	[17.5, 35]
Upper bound of preferred light intensity (coho salmon, large size)	I_{hpref}	PAR	[2.5, 4]
Upper bound of preferred light intensity (yellowtail amberjack)	I_{hpref}	PAR	[341.55, 805]
Upper bound of preferred light intensity (red seabream)	I_{hpref}	PAR	368
Steepness parameter of light reaction	I_{lsteep}	PAR	-20
Steepness parameter of light reaction	I_{hsteep}	PAR	1000
Lower bound of preferred temperature (coho salmon)	T_{lpref}	°C	8
Lower bound of preferred temperature (yellowtail amberjack)	T_{lpref}	°C	18
Lower bound of preferred temperature (red seabream)	T_{lpref}	°C	20
Upper bound of preferred temperature (coho salmon)	T_{hpref}	°C	9
Upper bound of preferred temperature (yellowtail amberjack)	T_{hpref}	°C	28
Upper bound of preferred temperature (red seabream)	T_{hpref}	°C	28
Steepness parameter of temperature reaction	T_{lsteep}	°C	-60
Steepness parameter of temperature reaction	T_{hsteep}	°C	80
Temperature at water surface at noon	T_{surf}	°C	8.3

E   Deep Reinforcement Learning

Proximal Policy Optimization   Proximal Policy Optimization (PPO) algorithm [27] improves upon vanilla policy gradient method by providing more stability and reliability during learning. It is a variant of Trust Region Policy Optimization (TRPO) algorithm [28] with significantly simpler implementation by optimizing a surrogate loss instead of KL divergence constraints. In our work, the clipped surrogate loss $L^{\text{CLIP}}(\theta)$ with respect to the current policy parameters θ is defined as follows:

$$L^{\text{CLIP}}(\theta) = \mathbb{E}_{s_t,a_t}\left[\min(l_t(\theta)\hat{A}_t, \text{clip}(l_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right]$$

$$l_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$$

π_θ and $\pi_{\theta_{\text{old}}}$ are current and old policies. $l_t(\theta)$ is the likelihood ratio of the action probability under current policy over that of old policy. As the new policy deviates from the old policy, $l_t(\theta)$ will deviate from 1 and be constrained by the interval $[1-\epsilon, 1+\epsilon]$. This means whenever excessive policy update happens the gradient will be set to zero and therefore prevents the new policy from being too different from previous one. $\hat{A}_t$ represents the generalized advantage estimation [29]. A minimum value is then chosen to obtain a lower bound of loss.

Since the parameters between the policy and value function are shared in networks, we further integrate the squared-error loss of value function ($L_t^{\text{VF}}(\theta)$) and an entropy term ($S[\pi_\theta](s_t)$) ensuring sufficient exploration to obtain the final objective loss:

$$L_t^{\text{CLIP+VF}+S}(\theta) = \mathbb{E}_{s_t,a_t}\left[L_t^{\text{CLIP}}(\theta) - cL_t^{\text{VF}}(\theta) + \beta S[\pi_\theta](s_t)\right]$$

where value function loss coefficient c is valued at 0.5 and entropy coefficient β is valued at 0.0005.

3

Parameters in Action and Reward Definition   The values of parameters used in action definition and reward function for PPO are listed in Table 3. The weight of each reward term is set according to the objective of each training. For example, we set the boundary avoidance weight to be the highest when the DRL controller was trained for the very first time because the primary goal of the training for the fish was to avoid the collision to the fish cage. When it is learnt, we increased the weights for the neighbor collision penalty and energy consumption penalty so that fish learn the smooth schooling in an environment as well as energy saving, which is an important benefit of being in a school [30].

Table 3: Parameters in action definition and reward function

Description	Parameter	Unit	Value
Sensing range (coho salmon and red seabream)	d_{sense}	BL	3
Sensing range (yellowtail amberjack)	d_{sense}	BL	2
Clamping angle of rotation about z-axis	θ^{zt}	Degree	10
Pitch angle threshold	θ^{rt}	Degree	53
Speed threshold	v^{st}	BL/second	0.3
Probability to start chase mode at each time step	p_a	N/A	0.005
Neighbor collision weight (pretraining)	w^{NC}	N/A	0.5
Neighbor collision weight (transfer learning)	w^{NC}	N/A	4
Boundary avoidance weight	w^{BD}	N/A	2
Neighbor interaction weight	w^{ND}	N/A	1.5
Rotation penalty weight (pretraining)	w^{r}	N/A	0.001
Rotation penalty weight (transfer learning)	w^{r}	N/A	0.5
Speed penalty weight (pretraining)	w^{s}	N/A	5
Speed penalty weight (transfer learning)	w^{s}	N/A	10
Chase reward weight for aggressors	w^{agg}	N/A	8
Escape penalty weight for targets	w^{tar}	N/A	1

PPO Training Configuration   We present the training hyperparameters of PPO in Table 4. Similar configurations were used for pretraining and transfer learning phases except the total number of steps when training coho salmon. This is primarily due to the variation in environment configurations, which makes retraining necessary. We also include an illustration of the learning curves of the three species in Fig. 1 to showcase the performance improvements over the entire training process.

SAC Training Configuration   We report the training hyperparameters of Soft Actor-Critic (SAC) algorithm [31] that we employed to train the fish agents aside from PPO in Table 5. Since we only compare the results of SAC and PPO in the default environment of coho salmon, the total numbers of time steps for SAC agents in pretraining and fine tuning phases are also four and one million.

Table 4: Hyperparameters used in training fish agents with PPO.

Parameter	Value
Value function loss coefficient for PPO c	0.5
Entropy coefficient for PPO β	0.0005
PPO clip threshold ϵ	0.2
Regularization parameter for GAE λ	0.95
Batch size	1024
Buffer size	10240
Learning rate	0.001
Num. of epochs	3
Discount factor γ	0.99
Time horizon	64
Total number of steps (pretraining, coho salmon)	4 million
Total number of steps (transfer learning, coho salmon)	1 million
Total number of steps (pretraining, yellowtail amberjack and red seabream)	2 million
Total number of steps (transfer learning, yellowtail amberjack and red seabream)	2 million

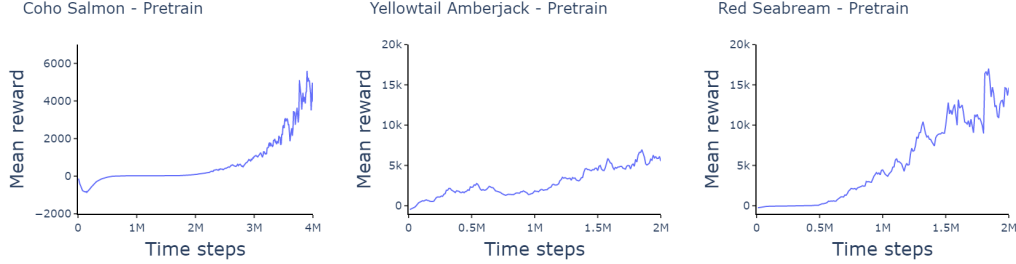
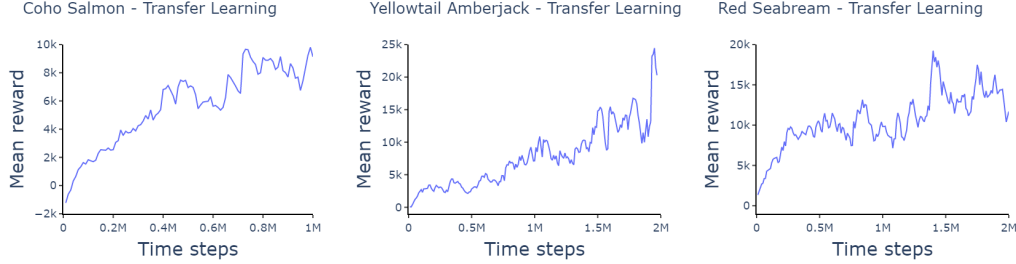
4

(a) Mean cumulative reward over all agents at each episode in pretraining phase. Left-to-right: coho salmon, yellowtail amberjack and red seabream

(b) Mean cumulative reward over all agents at each episode in transfer learning phase. Left-to-right: coho salmon, yellowtail amberjack and red seabream

Figure 1: Learning curves from PPO of the three fish species in pretraining and transfer learning phases in their corresponding default environments. Performance is measured by mean cumulative reward per episode smoothed by exponential moving average. Notably, coho salmon took four million time steps during pretraining followed by one million during transfer learning because of its larger quantity. Nevertheless the other two species were trained for two million steps in both phases.

F Physically Based Environment Simulation

Light Attenuation The simulation colors the water by first marching rays in fixed distance steps for every pixel on screen. The ray's direction is determined based on the pixel's location when transformed into world space. The current color of the pixel is then treated as incoming light, and becomes the input for the attenuation calculation using the Beer-Lambert Law shown in Eq. (1). I_0 represents the incoming light (or in this case the current associated pixel color) of the ray, a (or in this case $a_{r/g/b}$) is the attenuation coefficient of red, green or blue color in the vector $\mathbf{ac}$ which is determined by three components that we explain next, and d is the distance the ray of light has traveled through the water [32]. Then the simulation calculates the ambient light at the current point the ray is at. This is calculated by first summing up all the light that is heading towards the current

Table 5: Hyperparameters used in training coho salmon agents with SAC.

Parameter	Value
Target smoothing coefficient τ	0.005
Initial entropy coefficient α (pretraining)	0.3
Initial entropy coefficient α (fine tuning)	0
Average steps per update of policy	1000
Batch size	1024
Replay buffer size	102400
Learning rate (pretraining)	0.001
Learning rate (fine tuning)	0.0005
Discount factor γ	0.99
Time horizon	64
Total number of steps (pretraining)	4 million
Total number of steps (fine tuning)	1 million

5

point, and multiplying it by the dot product of the sunlight direction and ray direction. Next, the output of the ambient calculation is multiplied by the output of the attenuation calculation and is then added to a running total. Then the ray marches for a step in its set direction and the process repeats until it reaches its maximum distance. Once that is done, the running total is returned as the output color for the pixel associated with the given ray.

The attenuation coefficient vector consists of three parts: the attenuation of the water itself, the attenuation of chlorophyll, and the attenuation of the suspended sediments. Each part is a 3 dimensional vector representing the attenuation coefficients for red, green, and blue. The values of these coefficient vectors are determined based on data found in [32, 33, 34]. The coefficients of each are combined together, with the chlorophyll and sediments both multiplied by the concentration of their respective parameters using the model described in [32] and [35]. The formula is as follows:

$$\mathbf{ac} = \mathbf{ac}_{\mathrm{bw}} + co_{\mathrm{chl}}\mathbf{ac}_{\mathrm{chl}} + co_{\mathrm{sed}}\mathbf{ac}_{\mathrm{sed}} \tag{4}$$

where $\mathbf{ac}_{\mathrm{bw}}$ is the attenuation coefficient vector of the base water, co_{chl} is the concentration of chlorophyll, $\mathbf{ac}_{\mathrm{chl}}$ is the attenuation coefficient vector of chlorophyll, co_{sed} is the concentration of the sediments, and $\mathbf{ac}_{\mathrm{sed}}$ is the attenuation coefficient vector of sediments.

The factors such as the concentrations of sediments and chlorophyll are set by the simulation to random values from within user defined ranges. These ranges are based off values found in real world, like those laid out in [33] and [35]. The scene uses these factors to determine everything else about the environment. For example, the turbidity of the water is determined by the concentration of suspended sediments.

Light Scattering A light scattering system computes the light refraction based on the medium's properties such as the particles inside a fluid [36]. It uses a phase function, which can be obtained from a volume scattering function (VSF), to calculate the amount of light traveling through a medium and being scattered towards the viewer. We use the Kopelevich model [37] as the VSF to determine how different wavelengths of light are scattered underwater using large and small suspended particles. It can be formulated as the following equation [38]:

$$\mathrm{VSF} = \beta_w + v_s\beta_s\psi\left(\frac{500}{\lambda}\right)^{1.7} + v_l\beta_l\psi\left(\frac{500}{\lambda}\right)^{0.3} \tag{5}$$

where β_w, β_s and β_f denote the VSF of pure sea water, small and large particles respectively and their values can be selected from tabular values in [37] by performing visual inspection. v_s and v_l are the concentrations of small and large particles (in units of parts per million). ψ is the angle between the light and view directions, and λ is the wavelengths of red, green and blue. The resulting VSF is stored in a vector that is used to calculate the scattering coefficients for the three colors via numerical integration using the trapezoidal rule. Finally, the VSF vector is divided by the scattering coefficients to obtain the phase function for each of the colors, which is then used to multiply the output of the ray marching system to scatter the color based on the RGB value.

G Characteristics of fish school states

Data provided by the simulation are body length l_{BL}, position $\mathbf{x} = (x, y, z)$ of fish body center, speed v and its direction vector $\mathbf{d} = (d_x, d_y, d_z)$ ($|\mathbf{d}| = 1$) for each individual fish and each time slice. The velocity vector is $\mathbf{v} = v\mathbf{d}$, and the deviation from the center of the school $\mathbf{x}_{\mathrm{c}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$ is $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_{\mathrm{c}}$, where N denotes the number of fish in the school.

Order parameters According to [39, 40] two parameters, a polar order parameter P and a normalized angular momentum M of the school, are used to characterize state of school. The parameters P and M are given respectively as

$$P = \left|\frac{1}{N}\sum_{i=1}^{N}\mathbf{d}_i\right|, \quad M = \left|\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{r}_i \times \mathbf{v}_i}{|\mathbf{r}_i||\mathbf{v}_i|}\right| = \left|\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{r}_i \times \mathbf{d}_i}{|\mathbf{r}_i|}\right|, \tag{6}$$

in the same manner as in Ref.[40]. All fish move in the same direction in parallel when $P \simeq 1$ [39, 40], which cannot be observed in a fish cage. A large $M \simeq 1$ and small $P \simeq 0$ brings about the milling state, in which fish in a school rotate around an axis regularly [39, 40]. The smaller M and P indicate that the state is swarming state rather than milling state [39].

6

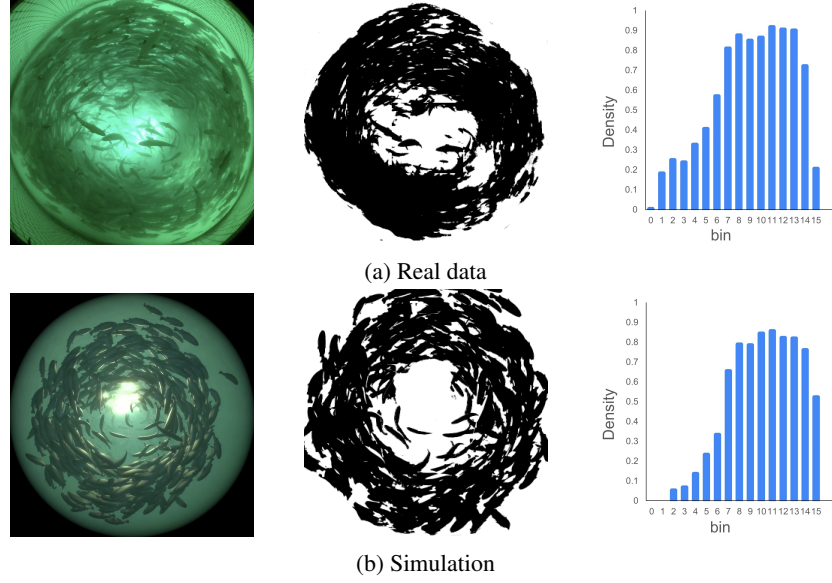
(a) Real data

(b) Simulation

Figure 2: Comparison between (a) real and (b) simulated obtained images. For both (a) and (b), left, middle, and right panels are respectively the fish-eye view picture, black and white images made from the left ones with some threshold, and density of black pixels for each bin obtained from the fish-eye images by dividing them into 16 annuli whose centers are the center of each picture.

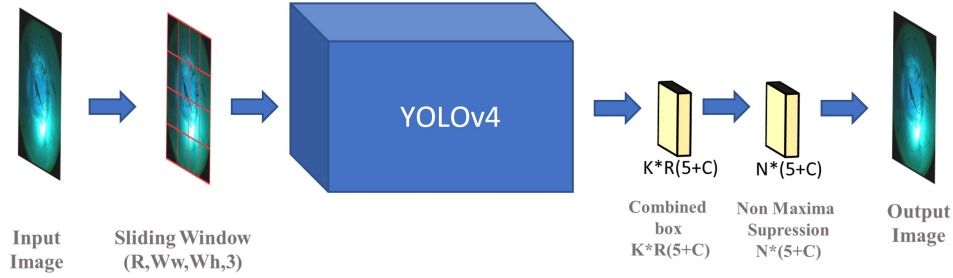



Figure 3: Fish detection model network architecture. This module is a region proposal network based on the YOLOv4 model. A sliding window is added as a first layer to crop sub-windows to magnify smaller fish in original images, which effectively avoids false negatives during smaller fish detection.

Similarity between pictures of real and the simulation   To see how the simulation results are consistent with the real data obtained from the field work, we compare the images obtained by both of them. These images are fish-eye lens view, and black and white images are made with some threshold from them. We compare the density of black pixels for each bin obtained by dividing each image into 16 annuli. As a result exhibited in Fig. 2, the simulation image seems to be consistent with the real one.

H   Fish Counting

Image Pre-Processing   The first module converts the input video to a sequence of images and applies denoising to produce the data format required by the fish detection module. The denoising process applies median blur to the images to reduce the effect of noise, resulting from caustics and sediment and other particles, often seen in underwater photography.

7

Fish Detection   Fig. 3 shows the network architecture of the fish detection model, based on YOLOv4 [41]. The input to the model is a sequence of images, and the output is a list of proposed bounding boxes (x, y, w, h), where (x, y) is the coordinate of the top left corner of the bounding box, and w and h are the width and height of the bounding box, respectively, with an associated class label and confidence score. We set the confidence score threshold to be 0.5 and any proposed regions below the threshold are discarded. In order to solve the issue of false negatives in the case of smaller fish, a sliding window is added as the first layer of the network. Each window crops a sub-image SI_i with a size of 416×416 from the original image I. The sliding window is shifted repeatedly horizontally and vertically by an amount equal to half the window size, producing a sub-image SI_i at each position that is fed to the YOLOv4 network. The produced number of sub-images, R, is computed using Eq. (7).

$$R = \left(\left| \frac{I_{\text{width}} - SI_{\text{width}}}{s} \right| + 1 \right) \times \left(\left| \frac{I_{\text{height}} - SI_{\text{height}}}{s} \right| + 1 \right), \qquad (7)$$

where SI_{width} and SI_{height} are the width and height of the sliding window, and I_{width} and I_{height} are the width and height of the original image, respectively. s is a step-size term and is set to be half of the sliding window size. This technique magnifies small fish in the original images. The computation time increases by a factor of R when processed serially; however, it helps to reduce the number of false negatives which is a major problem with the original YOLOv4. Note that we concatenate all the results from the cropped images into a single array, and apply Non-Maximum Suppression (NMS) to produce a single output, unlike the standard YOLO where the NMS is applied just after a single detection block.

Fish Counting   After the completion of the detection module, the estimation of fish count is tabulated for each frame, and the maximum value is taken as the final estimate of the number of fish in the cage. Despite its simplicity, we found the MAX method to give us the highest counting accuracy.

Training   We trained the model for 2000 epochs on a single Nvidia GeForce GTX 1080 Ti with an initial learning rate of 0.001. The training process completed in 5 hours. We then manually annotated 5 frames from 5 videos recorded at a real fish farm–a total of 25 images–and used them to fine-tune the model.

I   Field Data collection

We collected data on four species (two salmonids, yellowtail, red seabream) at marine fish cages. The salmonids cages are in the north, and the yellowtail and red seabream are in the south. Our goal was to measure the physical properties and dimensions of the environment and fish, to set the parameters of our simulation, and to record video of caged fish schooling behavior. In this appendix, we will discuss that data collection process and show our results.

Location   In March, May and July of 2021, we collected data from coho salmon and trout salmon cages on a northern fish farm (latitude 38.465, longitude 141.480). Both fish cages were cubes with a side length of 6.5m (Fig. 4(a)). Then, in June and November of 2021, we collected data from yellowtail and red seabream cages on a southern fish farm (latitude 34.215 and longitude 136.386). Both cages were cubes with side lengths of 3m (Fig. 4(b)).

Equipment   Fig. 5 shows the equipment used. Fig. 5a is circuit board used to measure light intensity. A Raspberry Pi is used to collect data from the TSL2571 light intensity sensor. To help it withstand the water pressure, the circuit board was placed in a clear smartphone case with the sensor positioned to avoid being obstructed. Additionally, the case was placed in a waterproof housing. After waterproofing, the sensor was attached alongside cameras to PVC pipes and used to measure light intensity at various depths underwater. Fig. 5b shows the device used to measure water temperature, the SK SATO SK1260. The sensor must be in direct contact with sea water, so it was lowered into the ocean. To film video of the fish schooling behaviors, we used KODAK 4KVR360 cameras. A camera placed at the center of the floor of the cage pointing up (Fig.5c) was fixed in place with four ropes. The cameras placed at the sides of the cage facing the center (Fig.5d) were fixed to PVC pipes and

8

| (a) In north | (b) In south | (c) Camera locations |

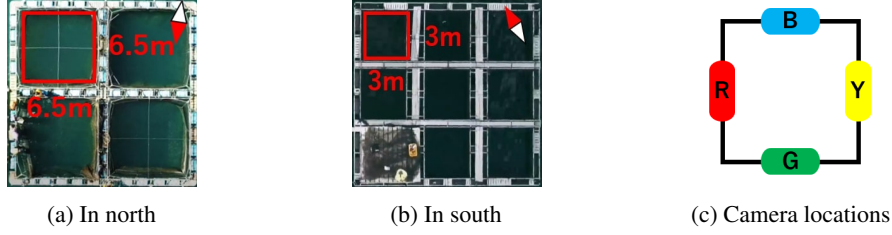



Figure 4: Fish firming cages and camera locations. (a) Salmonoids fish cages in north (coho salmon and trout salmon), (b) yellowtail and red sea bream fish cages in south, (c) camera locations in a cage.

| (a) | (b) | (c) |

| (d) | (e) | (f) |

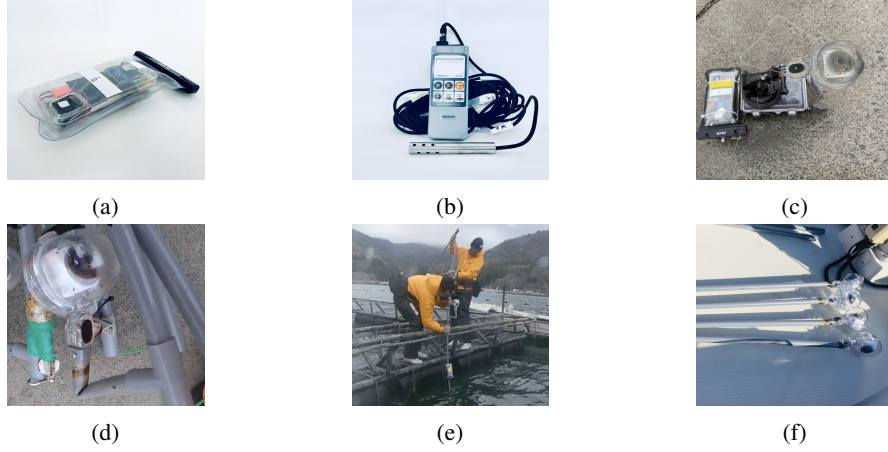



Figure 5: Equipment used at the fish cage. (a) light intensity sensor, (b) water temperature sensor, (c) center bottom camera, (d) side camera. A light intensity sensor is set behind the camera, (e) fixing PVC pipe, with cameras attached, to the cage, (f) cameras attached to a PVC pipe.

lowered to the target depth. When filming from multiple depths simultaneously, the cameras were fixed to long PVC pipes at 1m intervals, as in Fig.5f. The PVC pipe fixtures are shown in Fig.5e. By fixing the T-shaped part to the iron pipes of the cage, long video shoots are possible. The cameras are placed at R, G, B, and Y as shown in Fig. 4c.

Environment Data   Fig. 6a shows an example of the results of our light intensity measurements. The measurements were taken at the southern fish farm in November 2021, starting at 11:10 AM. The x-axis shows the number of minutes since the start, while the y-axis is the light intensity (Lux). Light blue shows the readings at a depth of 5m (center bottom), red shows readings at 3m (side), yellow shows readings at 2m (side), and black shows readings at 1m (side). It can be seen that light has lower intensity at greater depths. Sudden noisy drops in brightness are likely the result of fish blocking light from reaching the sensor. Fig. 6b shows the temperature at each depth. The top green line is data from the southern farm in June. The rest are data from the north farm, recorded in (from bottom to top) March, May, and July. Although there are seasons without any temperature variation, at other times the temperature can differ by 3-4 degrees. During the period from May to July, when water temperature exceeds 10 degrees, the chlorophyll content of the water increases, lowering the clarity of the water.

Body Measurements and Fish Count   The number of fish in a fish farm is unclear. Fish are purchased by weight as juveniles. This weight is divided by an average weight to give an approximate count, but this can result in up to 50% error. A fish's value on the market is determined by weight. Therefore, fish farmers try to monitor the state of their fish (length, height, width, and weight) in order to optimize their feeding, but the process is dependent on farmers' intuition. The only way to accurately measure a fish is by hand after removing it from the water[4]. To do so, fish are scooped

---

[4]These measurement methods, while designed to avoid injuring the fish, are neither completely safe nor particularly accurate. There is a demand for an alternative method to obtain these data from video data

9

(a) Light intensity   (b) Water temperature

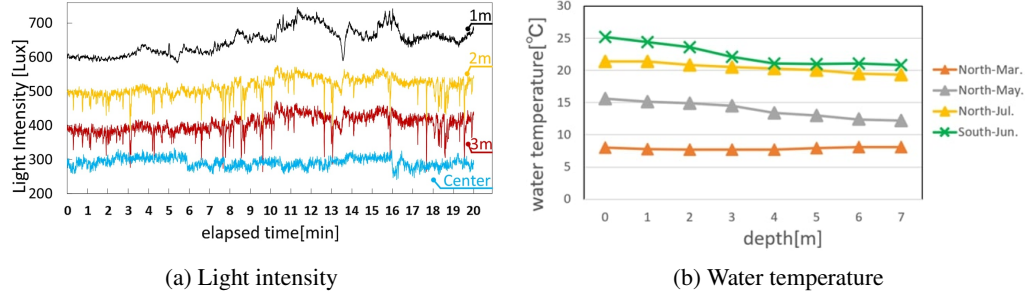



Figure 6: Environment data: (a) November 2021, Northern farm light intensity. From bottom to top: center bottom (5m), 3m, 2m, 1m. (b) Water temperature for each depth.

(a) Scooping fish   (b) Fish body measurement   (c) Fish counting

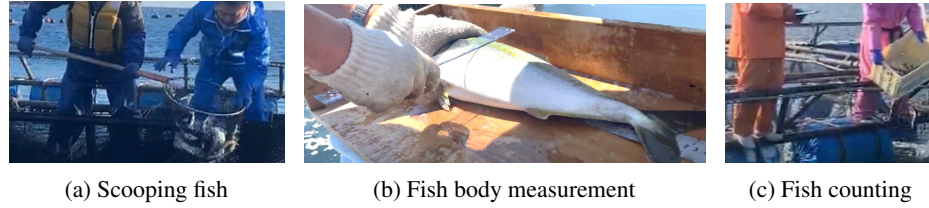



Figure 7: Fish body measurement and counting process

up by a net a few fish at a time; when the fish are released again, they are counted by eye. When populations are small this process can be performed for all fish, but it is not realistic for more than 1000 fish. For such cases, the process is performed for approximately 10% of the fish and the result is multiplied by 10, giving an inexact value. Fig. 7c shows the current fish counting process.

Fish body measurements are even more difficult. To measure a fish, the fish is removed from the water and anesthetized to prevent injury, and measured by hand with a tape measure (see Fig. 7). Fig. 8 shows the body lengths of 10 randomly selected fish. Fig. 8a shows the results for salmon (coho salmon and trout salmon), and Fig. 8b shows the results for yellowtail and red seabream. The x-axis shows the month the measurement was performed. The lines show the range of body lengths, and the symbol is the average length (coho salmon: gray, trout salmon: orange, yellowtail: yellow, red seabream: red). Salmon farming begins in December, and they are shipped to market in July. The graph shows that there exists large variation in salmon body length, which gets smaller as July approaches. The shipping date for yellowtail and red seabream is not fixed; they are shipped once they reach a set size.

Video Recording   Videos at the fish farm were filmed with KODAK 4KVR360 cameras. One camera was placed at the bottom center of the cage, facing the ocean surface. The others were attached to PVC pipes placed at the center of each side at 1m intervals. The footage captured in March, May, and July 2021 on the northern farm and June 2021 on the southern farm was filmed in the KODAK 4KVR360 235 degree dome mode. The November 2021 footage on the southern farm

(a)   (b)

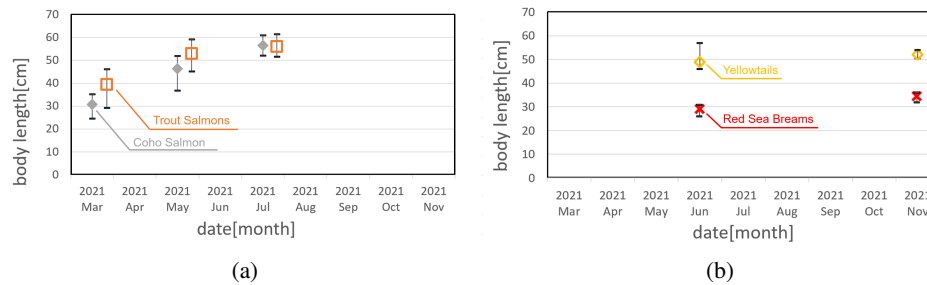



Figure 8: Body length: (a) Salmon body length. Orange: trout salmon, gray: coho salmon. (b) Yellowtail and red seabream body length. Red marks: Red seabream, yellow marks: yellowtails.

10

(a) Coho salmons at a depth of 3m in April, 2022 at north fish cage.

(b) Yellowtail amberjack at a depth of 3m in November, 2021 at south fish cage.

(c) Red seabreams at a depth of 2m in November, 2021 at south fish cage.

Figure 9: Examples of recorded video shots of (a) coho salmon, (b) yellowtail amberjack, and (c) red seabreams. For each fish species, from left to right panels show camera views located at R, G, B, and Y respectively.

was filmed in the KODAK 4KVR360 155 degree front mode. The coho salmon, yellowtail, and red seabream footage is show in Fig. 9. The cameras are placed at R, G, B, Y in Fig.4c. The salmon and yellowtail footage was taken at a depth of 3m, while the red seabream footage was taken at a depth of 2m. The footage at these depths contained the largest number of fish for each species, showing that the preferred depth differs per species.

References

[1] William Wallace Reynolds and Martha Elizabeth Casterlin. Behavioral thermoregulation and the "final preferendum" paradigm. *American zoologist*, 19(1):211–224, 1979.

[2] Arne J Jensen, Bjørn O Johnsen, and Laila Saksgård. Temperature requirements in atlantic salmon (salmo salar), brown trout (salmo trutta), and arctic char (salvelinus alpinus) from hatching to initial feeding compared with geographic distribution. *Canadian Journal of Fisheries and Aquatic Sciences*, 46(5):786–789, 1989.

[3] J Kohbara, I Hidaka, F Matsuoka, T Osada, K Furukawa, M Yamashita, and M Tabata. Self-feeding behavior of yellowtail, seriola quinqueradiata, in net cages: diel and seasonal patterns and influences of environmental factors. *Aquaculture*, 220(1-4):581–594, 2003.

[4] Tomoki Honryo, Michio Kurata, Dario Sandval, Saki Yamao, Amado Cano, and Yoshifumi Sawada. Effect of water temperature and light intensity on swim bladder inflation and growth of red sea bream pagrus major larvae. *Fisheries science*, 84(3):553–562, 2018.

[5] I Huse and JC Holm. Vertical distribution of atlantic salmon (salmo salar) as a function of illumination. *Journal of Fish Biology*, 43:147–156, 1993.

11

[6] Anders Fernö, Ingvar Huse, Jon-Erik Juell, and Åsmund Bjordal. Vertical distribution of atlantic salmon (salmo solar l.) in net pens: trade-off between surface light avoidance and food attraction. *Aquaculture*, 132(3-4):285–296, 1995.

[7] Jon-Erik Juell. The behaviour of atlantic salmon in relation to efficient cage-rearing. *Reviews in fish biology and fisheries*, 5(3):320–335, 1995.

[8] Frode Oppedal, Tim Dempster, and Lars H Stien. Environmental drivers of atlantic salmon behaviour in sea-cages: a review. *Aquaculture*, 311(1-4):1–18, 2011.

[9] Martin Føre, Tim Dempster, Jo Arve Alfredsen, Vegar Johansen, and David Johansson. Modelling of atlantic salmon (salmo salar l.) behaviour in sea-cages: A lagrangian approach. *Aquaculture*, 288(3-4):196–204, 2009.

[10] James E Herbert-Read, Andrea Perna, Richard P Mann, Timothy M Schaerf, David JT Sumpter, and Ashley JW Ward. Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences*, 108(46):18726–18731, 2011.

[11] Lucy Odling-Smee and Victoria A Braithwaite. The role of learning in fish orientation. *Fish and Fisheries*, 4(3):235–246, 2003.

[12] Alisha A Brown, Marcia L Spetch, and Peter L Hurd. Growing in circles: Rearing environment alters spatial navigation in fish. *Psychological Science*, 18(7):569–573, 2007.

[13] Horst Bleckmann. 3-d-orientation with the octavolateralis system. *Journal of Physiology-Paris*, 98(1-3):53–65, 2004.

[14] KF Cubitt, S Churchill, D Rowsell, DA Scruton, RS McKinley, et al. 3-dimensional positioning of salmon in commercial sea cages: assessment of a tool for monitoring behaviour. pages 25–33, 2003.

[15] Jens Krause, Jean-Guy J Godin, and David Brown. Phenotypic variability within and between fish shoals. *Ecology*, 77(5):1586–1591, 1996.

[16] Malthe Hvas, Ole Folkedal, and Frode Oppedal. Fish welfare in offshore salmon aquaculture. *Reviews in Aquaculture*, 13(2):836–852, 2021.

[17] Ole Folkedal, Lars Helge Stien, Jonatan Nilsson, Thomas Torgersen, Jan Erik Fosseidengen, and Frode Oppedal. Sea caged atlantic salmon display size-dependent swimming depth. *Aquatic Living Resources*, 25(2):143–149, 2012.

[18] C.A.G. Wiersma and K. Ikeda. Interneurons commanding swimmeret movements in the crayfish, procambarus clarki (girard). *Comparative Biochemistry and Physiology*, 12(4):509–525, 1964.

[19] Henri Korn and Donald S Faber. The mauthner cell half a century later: a neurobiological model for decision-making? *Neuron*, 47(1):13—28, July 2005.

[20] Thomas H. Miller, Katie Clements, Sungwoo Ahn, Choongseok Park, Eoon Hye Ji, and Fadi A. Issa. Social status–dependent shift in neural circuit activation affects decision making. *Journal of Neuroscience*, 37(8):2137–2148, 2017.

[21] Charles L. Gallegos and Moore. Factors contributing to water-column light attenuation. *Chesapeake Bay submerged aquatic vegetation water quality and habitat-based requirements and restoration targets: a second technical synthesis*, pages 35 – 55, 2000.

[22] Annick Bricaud, André Morel, Marcel Babin, Karima Allali, and Hervé. Claustre. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research*, 1998.

[23] Michael Solonenko and Curtis D. Mobley. Inherent optical properties of jerlov water types. *Applied optics*, 54 17:5392–401, 2015.

12

[24] Kevin R Stuart and Mark Drawbridge. The effect of light intensity and green water on survival and growth of cultured larval california yellowtail (seriola lalandi). *Aquaculture*, 321(1-2):152–156, 2011.

[25] John P Abraham, M Baringer, NL Bindoff, T Boyer, LJ Cheng, JA Church, JL Conroy, CM Domingues, JT Fasullo, J Gilson, et al. A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Reviews of Geophysics*, 51(3):450–483, 2013.

[26] Yoshimi Kawai and Akiyoshi Wada. Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: A review. *Journal of oceanography*, 63(5):721–744, 2007.

[27] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[28] John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International conference on machine learning*, pages 1889–1897. PMLR, 2015.

[29] John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. High-dimensional continuous control using generalized advantage estimation. *arXiv preprint arXiv:1506.02438*, 2015.

[30] Stefano Marras, Shaun S Killen, Jan Lindström, David J McKenzie, John F Steffensen, and Paolo Domenici. Fish swimming in schools save energy regardless of their spatial position. *Behavioral ecology and sociobiology*, 69(2):219–226, 2015.

[31] Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018.

[32] Charles L Gallegos and KA Moore. Factors contributing to water-column light attenuation. *Chesapeake bay submerged aquatic vegetation water quality and habitat-based requirements and restoration targets: A second technical synthesis*, 2000.

[33] Annick Bricaud, André Morel, Marcel Babin, Karima Allali, and Hervé Claustre. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research: Oceans*, 103(C13):31033–31044, 1998.

[34] Derya Akkaynak, Tali Treibitz, Tom Shlesinger, Yossi Loya, Raz Tamir, and David Iluz. What is the space of attenuation coefficients in underwater computer vision? In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4931–4940, 2017.

[35] Michael G Solonenko and Curtis D Mobley. Inherent optical properties of jerlov water types. *Applied optics*, 54(17):5392–5401, 2015.

[36] Nathan Hoobler. Fast, flexible, physically-based volumetric light scattering. In *Game Developers Conference*, page 608, 2016.

[37] OV Kopelevich. Small-parameter model of optical properties of sea water. *Ocean optics*, 1:208–234, 1983.

[38] Curtis D Mobley. In *Light and Water: Radiative Transfer in Natural Waters*. Academic, 1994.

[39] Johann Delcourt and Pascal Poncin. Shoals and schools: back to the heuristic definitions and quantitative references. *Reviews in Fish Biology and Fisheries*, 22(3):595–619, 2012.

[40] Daniel S Calovi, Ugo Lopez, Sandrine Ngo, Clément Sire, Hugues Chaté, and Guy Theraulaz. Swarming, schooling, milling: phase diagram of a data-driven fish school model. *New Journal of Physics*, 16(1):015026, jan 2014.

[41] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.

13