



# Automated computer vision based individual salmon (*Salmo salar*) breathing rate estimation (SaBRE) for improved state observability

Espen Berntzen Høgstedt <sup>a,b,\*</sup>, Christian Schellewald <sup>c</sup>, Rudolf Mester <sup>b</sup>, Annette Stahl <sup>a</sup>

<sup>a</sup> Department of Engineering Cybernetics, Norwegian University of Science and Technology, O. S. Bragstads Plass 2D, 7034 Trondheim, Norway

<sup>b</sup> Department of Computer Science, Norwegian University of Science and Technology, IT-Bygget, Almen, Sem Sælands vei 9, 7034 Trondheim, Norway

<sup>c</sup> SINTEF Ocean, Otto Nielsens Veg 10, 7052 Trondheim, Norway

## ARTICLE INFO

### Keywords:

Computer vision  
Deep learning  
Respiration frequency  
Atlantic salmon  
Fish welfare  
Aquaculture

## ABSTRACT

Salmon farming plays an important role in the Norwegian food industry, supplying a large portion of the world's salmon. Despite its economic importance, salmon farming faces unique biological challenges impacting the health and welfare of the fish. These factors limit the growth of salmon and the full economic potential of salmon farms, highlighting the need for advanced techniques in aquaculture to enhance productivity and sustainability. A key technical challenge is the unavailability of effective monitoring tools, leading to greater difficulties assessing the condition of salmon in water compared to the assessment of the animal stock on land-based farms. To improve the observability of salmon farms, we designed and created a computer-vision based approach for salmon breathing rate estimation (SaBRE), which allows the automatic monitoring of the respiration frequency of each individual salmon in a group of fish, seamlessly covering the entire workflow from video-stream input to the final data output (end-to-end). We thoroughly evaluated the capabilities of our method in two ways. Firstly, we performed a quantitative analysis of the constituent modules of SaBRE, revealing that all modules were highly accurate, including a salmon re-identification module that achieved an accuracy of 99.51 %. Secondly, we analyzed data from a salmon experiment with SaBRE, demonstrating that our algorithm provides high-quality respiration frequency information that can be compared with other types of experimental data to infer biological relationships. For the fish in our experiment, we observed that the ranking of the respiration frequencies in individual salmon remains relatively unchanged both in the short term and over longer periods (i.e., a salmon with a high breathing rate consistently remains among those with the highest rates, regardless of changes in the environment). Furthermore, a significant negative correlation (Pearson correlation coefficients with  $r$  values between  $-0.61$  and  $-0.90$  and  $p$  values below 0.01) was observed between our algorithm's estimated respiration frequency and the dissolved oxygen ( $dO_2$ ) content in the water. In addition, the average breathing rate of the salmon was observed to increase in response to incidents potentially causing stress.

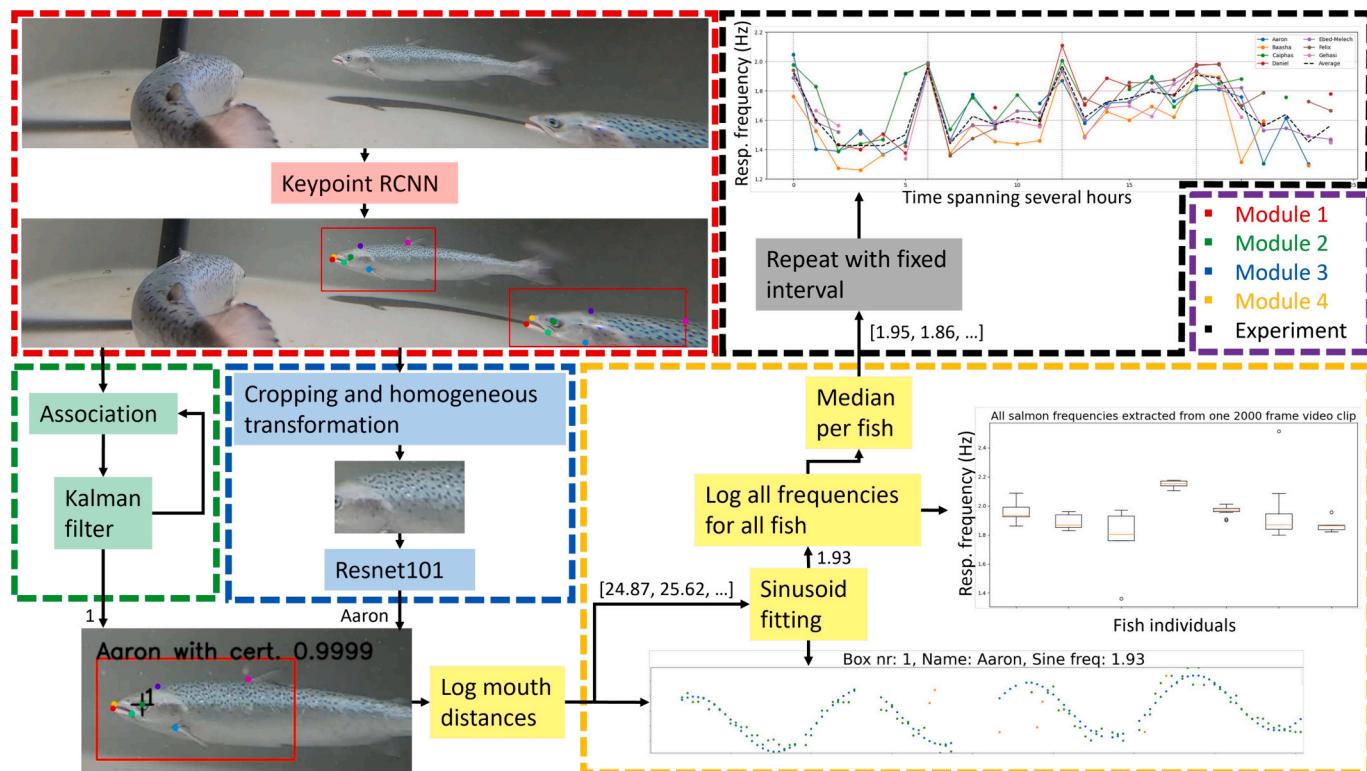
## 1. Introduction

Finfish aquaculture is a rapidly developing industry (Fry et al., 2018), emerging as a sustainable and effective means of feeding an expanding population (Torrisen et al., 2011). Today, sea cages in the Norwegian salmon industry typically measure 157 m in circumference and extend to depths of 50 m (Føre et al., 2018). One farming locality consists of six to ten circular sea cages, each capable of holding up to 200,000 salmon (Øvrebø et al., 2022). The farming of salmon, however, comes with a series of challenges. High temperatures cause temperature-induced cessation of voluntary feed intake as well as inefficient

osmoregulatory, liver and renal function (Wade et al., 2019), while low dissolved oxygen ( $dO_2$ ) content (caused by poor water exchange (Oldham et al., 2018) and high biomass) and stress reduce the growth and welfare of the fish (McCormick et al., 1998; Olsen et al., 2002; Pandurangan et al., 2014; Pettersen et al., 2013; Solstrom et al., 2018). To combat these adverse effects, the aquaculture professionals operating the cages are dependent on high-quality information describing the health and well-being of more than a million fish with subtle and difficult-to-interpret behavioral cues. This requires effective measurement strategies of various farm states, including the easier-to-measure concentration of  $dO_2$  and temperature or the difficult-to-measure

\* Corresponding author at: Department of Computer Science, Norwegian University of Science and Technology, IT-Bygget, Almen, Sem Sælands vei 9, 7034 Trondheim, Norway.

E-mail address: [espen.b.hogstedt@ntnu.no](mailto:espen.b.hogstedt@ntnu.no) (E.B. Høgstedt).



**Fig. 1.** The initial module of SaBRE predicts bounding boxes and keypoints for all salmon in a frame. The upper jaw is shown as orange, the lower jaw as red, the eye as green, the root of the jaw as turquoise, the root of the pectoral fin as blue, the cranial root as purple and the anterior root of the dorsal fin as pink. From the salmon detection and keypoint prediction information, module 2 tracks individual salmon between frames by a Kalman filter approach, while module 3 identifies the fish individual in each bounding box. Module 2 and module 3 are then cooperating to generate time series of the jaw gape evolution of a single fish instance over L frames, before this time series is fitted to a sinusoid by combining Levenberg-Marquardt function fitting and random sample consensus (RANSAC). All sinusoid frequencies over a specific frame range are sorted into which fish individual they arose from, and the per-fish respiration frequency median is calculated to provide the momentary salmon respiration frequency for each fish. Finally, evaluating momentary per-fish respiration frequency with a specific time interval allows evaluating the evolution of salmon respiration frequency. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

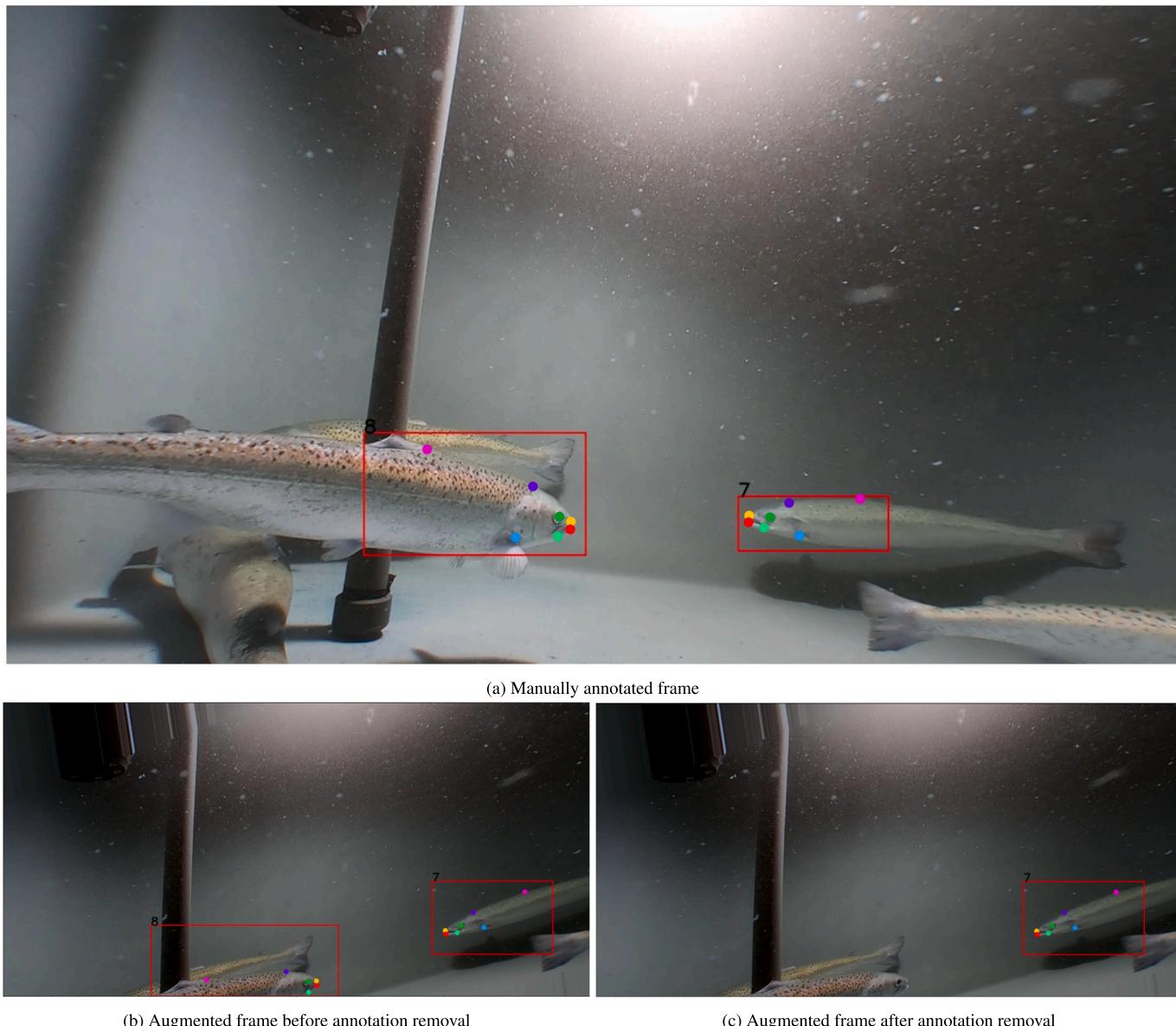
stress levels of salmon. Towards an implementation of more sophisticated, efficient, and welfare-oriented salmon aquaculture practices, continuous and accurate monitoring of the farm states is highly desired.

The difficulty of measuring farm states varies greatly. Temperature is easy to measure, and accurate electronic devices are available at a low cost (Parra et al., 2018). For other farm states, available monitoring methods are often a compromise between accuracy and cost for farmers. dO<sub>2</sub> content can be measured relatively precisely at single locations, but due to the complex dynamics of farm sites, a large number of dO<sub>2</sub> sensors at different locations are required to obtain a sufficient understanding of the dO<sub>2</sub> distribution in the salmon cages (Solstrom et al., 2018). Furthermore, measuring dO<sub>2</sub> content is an intrinsically difficult task, and state-of-the-art methods are still prone to drift and instability (Wei et al., 2019). Although mathematical modeling (Alver et al., 2022) and intelligent sensors (Wei et al., 2019) are making progress in reducing the shortcomings of the available sensor technology, dissolved oxygen measurement is still an area where technical improvements can have a significant impact. For the measurement of stress, sampling and subsequent chemical analysis of blood or organs of fish is the current gold standard (Sopinka et al., 2016), which is expensive, labor-intensive, non-continuous and infeasible to perform at individual-fish level/niveau within commercial farms. There have been some attempts at measuring stress by environmental analysis (Ellis et al., 2004), which solves several of the shortcomings of blood sampling, but they are still incapable of tracking the stress responses of individual salmon. In modern aquaculture facilities, a significant part of state inference is still done by visual interpretation from human observers, where farmers are reviewing swimming behavior (Martins et al., 2012; Nilsson et al.,

2022), or external appearance (Noble et al., 2018) to infer about salmon welfare and stress. Although this strategy is informative, it is labor intensive and requires expert knowledge. Additionally, even with excessive human labor, such an approach is not applicable to state monitoring of single fish over extended periods, both due to the number of salmon in commercial farms, and due to the limited ability of humans to re-identify salmon individuals.

In recent years, breakthroughs in the field of computer vision have enabled the automated detection and analysis of visible physiological parameters and biometric features, paving the way for cheap, automated and continuous state monitoring of individual fish. To allow farmers or automated control systems to respond quickly to adverse situations, the targeted external salmon traits should be informative of a diverse range of salmon states, as well as manifesting quickly externally as the environmental conditions change. One such trait is swimming behavior (Nygård et al., 2022), which has been explored for automatic inference on feeding activity (Liu et al., 2014), hypoxia response (Israeli and Kimmel, 1996) and welfare (Pinkiewicz et al., 2011). Another far less explored external trait for automated fish state monitoring is respiration frequency, which is easily discernible, quickly adapting to altered conditions (Ballintijn, 1972), and tightly linked to fish states such as stress (Aota et al., 1990; Kinhead and Perry, 1991; Randall and Taylor, 1991; Wendelaar, 1997), environmental dO<sub>2</sub> concentration (Remen, 2012) and gill disease (Bowden et al., 2022).

The exploration of respiratory frequency of fish through computer vision remains an underrepresented area in research. An initial study from 2014 conducted by Zheng et al. (2014), monitored the respiration rate of medaka (*Oryzias latipes*) fish in small, individual containers by



**Fig. 2.** Visualization of three steps in the creation of the salmon keypoint dataset. The upper jaw is shown as orange, the lower jaw as red, the eye as green, the root of the jaw as turquoise, the root of the pectoral fin as blue, the cranial root as purple and the anterior root of the dorsal fin as pink. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observing changes in the number of red gill pixels in images taken by a camera mounted at the bottom of a tank. A paper by Kämmer et al. (2023) focused on the respiratory patterns of zebrafish (*Danio rerio*) embryos and larvae. The employed method included anesthetizing the fish, putting them in specialized gel blocks, filming them individually under a microscope, and tracking the changes in grayscale pixel values between frames. A recent study (Schellewald et al., 2024) introduced a method to measure the mouth-opening frequency of salmon from underwater videos, utilizing optical-flow techniques.

The few available methods for fish respiration monitoring all require highly specialized setups with separated fish individuals, making their approaches inapplicable to individual fish monitoring in larger salmon groups, and monitoring in less controlled settings. By employing methods from the emerging field of deep learning, it is possible to construct far more powerful and generally applicable algorithms. The core idea behind these methods is to iteratively update the weights of a huge general function approximator (deep artificial neural network) so that it can associate input images with desired output annotations as

closely as possible. Current state-of-the-art networks typically consist of a backbone of convolutional layers or attention blocks (Maurício et al., 2023), which extract compact and informative features of the input image, and a head that combines these features to approximate the annotated output. The head can be constructed in a way that allows the network to specify the locations of objects in an image (Redmon et al., 2016), or specify the class (He et al., 2016), segmentation masks or keypoints (He et al., 2017) of an input object. Some networks are constructed in a way that solves several of these problems simultaneously, such as the Keypoint RCNN (He et al., 2017), which predicts object location, object class and object keypoints in parallel from the same backbone features.

A significant shortcoming of earlier respiration frequency monitoring methods for fish is the lack of long-term fish recognition, which disallows automated monitoring of the ventilation of a single fish over hours and days if the fish is reared with other fish. As demonstrated in earlier works (Cisar et al., 2021; Stien et al., 2017), the melanin pattern on salmon can distinguish individuals, at least in smaller groups of fish.

Deep neural networks can learn to recognize such distinguishing features to precisely re-identify individual objects. For smaller groups, re-identification can be performed by treating each object's identity as a class. For larger groups of fish, a deep network can learn to generate a compact feature representation that is equal for equal identities, and invariant to other image changes such as object pose, object illumination and the distance between the camera and the object (Gómez-Vargas et al., 2023; Mathisen et al., 2020; Zhou et al., 2022).

Driven by the limited research on computer vision-based respiration analysis methods for fish and the opportunity to gain insights into animal welfare through respiration data, we designed and implemented a new system capable of conducting a fully automated salmon breathing rate estimation (SaBRE). The method is specifically suitable for small groups of salmon in smaller tanks, where it continuously monitors the respiration frequency of each individual salmon. This is achieved by a pipeline consisting of four modules (Fig. 1): salmon detection and pose (anatomical landmark) estimation using the Keypoint RCNN (He et al., 2017) architecture, salmon re-identification (including a method for semi-automated labeling that enables the construction of a large individual salmon dataset by the expenditure of limited human effort) using the ResNet 101 architecture (He et al., 2016), salmon tracking over the short term by using a Kalman filter (Kalman, 1960) and a variant of the Jonker-Volgenant algorithm (Crouse, 2016), and mouth frequency analysis by fitting a sinusoid to mouth gape time series by combining the Levenberg-Marquardt algorithm (Lourakis, 2005) with a RANSAC loop (Raguram et al., 2013). Subsequently, we examined the capabilities of the method for state observation by using it to analyze data from a salmon experiment in which 7 fish in a 1000-l tank were subjected to disturbances and oxygen depletion.

Salmon respiration frequency is typically defined as the number of opercular oscillations every minute. In the work done by Kämmer et al. (2023), this definition was expanded to also include the motion of jaws, which seems reasonable when considering the intimate link between gill opening and mouth opening during buccal pumping (Taylor, 2011). Therefore, we adopt in the present paper the same assumption as Kämmer, namely that both the jaw gape motion and the gill pulsation approximately describe the respiratory pattern of fish. As both jaw-gape motion and salmon dot pattern are visible from the side of the fish, this assumption facilitates the identification of salmon and the determination of respiration frequency from the same images.

## 2. Materials and methods

### 2.1. Module 1: salmon and keypoint detection

The first module of SaBRE aims to detect all salmon in a single video frame and locate anatomical landmarks on these fish. This was accomplished by employing a PyTorch (Paszke et al., 2019) Keypoint RCNN model (He et al., 2017), pretrained on Imagenet (Deng et al., 2009), and fine-tuned on a custom salmon keypoint dataset.

When labeling the dataset, bounding boxes were placed such that all of the salmon rostral to the caudal root of the dorsal fin were inside the box, while the keypoints were placed at the eye, the root of the pectoral fin, the cranial root of the dorsal fin, the tip of the premaxilla, the tip of the kype, the intersection between the mandible and the maxilla, and the dorsalmost edge of the skull (Fig. 2).

The salmon keypoint dataset was constructed by first manually annotating 200 frames with keypoints and bounding boxes for each of the fish visible in the images, and then inflating the dataset by a factor of 30 by subjecting the images to random contrast adjustment, brightness adjustment, rotation, four-point perspective transformation, shearing, scaling and cropping by the use of the python Albumentations (Buslaev et al., 2020) library. During the augmentation process, all incomplete labels were removed (Fig. 2), and in the case of no remaining labels in an

image, a new augmented image was generated.

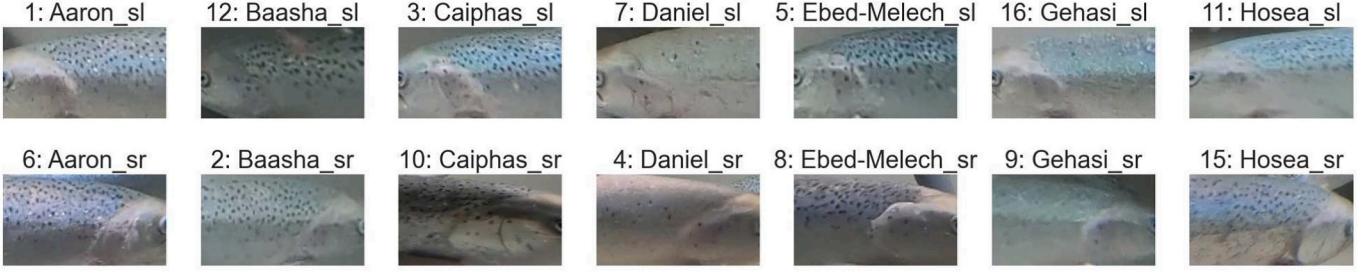
To fine-tune the Keypoint RCNN network, the keypoint and bounding box heads (last network layers) of the pretrained model were replaced with randomly initialized heads matching the dimensions of our problem, before the complete parameter set of the model was trained on a computer with 2 T4 GPUs and 1 Intel Xeon CPU for 13 epochs on the salmon keypoint dataset. The momentum was set to 0.9, the initial learning rate to 0.005, the learning rate decay to a factor of 10 every 3 epochs and the regularization to 0.0005. This yielded a total training time of eight to nine hours on the used hardware. Finally, the bounding box set was refined by the PyTorch (Paszke et al., 2019) non maximum suppression method during inference, with an overlap threshold of 0.1.

### 2.2. Module 2: salmon tracking

The salmon tracking (see [algorithm 1](#) at the end of this subsection) was performed by keeping a set of updated Kalman filters (Kalman, 1960)  $\mathcal{K}$ , where each filter  $\mathcal{K}[i]$  in this set contained information concerning one of the M salmon visible in the camera field of view. The  $\mathcal{K}[i]$  data objects contained both Kalman state variables (position  $\mathcal{K}[i]_{loc} \in \mathbb{R}^2$  and velocity  $\mathcal{K}[i]_{vel} \in \mathbb{R}^2$ ), and information necessary for determining individual respiration frequency in consecutive SaBRE modules. This information included the pose of the mouth of the salmon (given the location of the predicted jaw keypoints), and salmon identity class (i.e. the identity information of the salmon as discussed in [subsections 2.3 and 2.4](#)) for each frame. As a salmon eye center is a robust single-point representation of the location of a salmon, as well as being easily detectable by the Keypoint RCNN model of module 1 ([Subsection 2.1](#)), this anatomical landmark was used as the location anchor ( $\mathcal{K}[i]_{loc}$ ) when tracking the fish. The following three steps are executed for every new frame to update the location estimate for each tracked salmon. Square brackets ([ ]) are used to index set or vector positions (i.e.  $\mathcal{K}[i]$  refers to the  $i^{th}$  data object in the Kalman filter set).

1. The next state ( $\mathcal{K}[i]_{loc}$  and  $\mathcal{K}[i]_{vel}$ ) of each salmon tracker is predicted by the `predict()` method using a simple linear two-dimensional constant velocity model.
2. Trackers ( $\mathcal{K}[i] \in \mathcal{K}$ ) and eye detections from the Keypoint RCNN model ( $e \in E$ , with  $E \in \mathbb{R}^{N \times 2}$ ) are matched by minimizing the sum of the Euclidean distances between eye and tracker assignments. These matches are calculated by a variant of the `JonkerVolgenant()` method (Crouse, 2016), which accepts an Euclidean distance cost matrix ( $C \in \mathbb{R}^{M \times N}$ ) describing the distance between all detected eye positions ( $E$ ) and all tracker locations ( $\mathcal{K}[i]_{loc}$ ) in the tracker set ( $\mathcal{K}$ ), and outputs two index vectors ( $m_E \in \mathbb{N}^S$  and  $m_K \in \mathbb{N}^S$ ) specifying the indices of  $E$  and  $\mathcal{K}$  that are matches (i.e., the tracker  $\mathcal{K}[m_K[i]]$  is associated to the eye  $E[m_E[i]]$ ).
3. The predicted tracker states (see point 1. in this list) are updated by the `update()` method, which accepts a tracker  $\mathcal{K}[i]$  and its assigned eye detection  $e$  (point 2. in this list), and returns an updated  $\mathcal{K}[i]$  that contains Kalman states that have been modified as a weighted sum between the original states of  $\mathcal{K}[i]$  and the assigned measurement.

The used tracking method is explained in [algorithm 1](#) (using 1-indexing, i.e. the first element of a data structure is at index 1), with  $\mathcal{V}$  being an ordered set of image tensors, `incrUnmatched() :  $\mathcal{K}[i] \rightarrow \mathcal{K}[i]$`  increments a flag in a Kalman filter which tracks the time span the filter has not received a measurement, `unmatched() :  $\mathcal{K}[i] \rightarrow \mathbb{R}^1$`  returns the value of this flag,  $k$  sets the threshold for how large the `unmatched` value can grow before the corresponding tracker is inserted into the tracker *not active* set ( $\mathcal{K}_{log}$ ) and `newTracker() :  $\mathbb{R}^2 \rightarrow \mathcal{K}[i]$`  initializes a new tracker with zero velocity and position at the eye of the salmon it starts to track.



**Fig. 3.** Examples of operculum and anterior trunk (OAT) images for all 7 fish individuals in the tank. There are 14 classes, one for each side of the 7 fish in the tank. The displayed fish in these images have been automatically located and transformed by SaBRE modules, while the ID has been specified manually. The images are in full resolution.

**Algorithm 1.** Tracking of individual salmon in the camera field of view

task of annotating fish images to that of annotating trackers, each containing up to several hundred images, significantly improving the speed and precision of the annotations.

---

```

Require:  $\mathcal{V}, k$ 
 $\mathcal{K} \leftarrow \emptyset$ 
 $\mathcal{K}_{log} \leftarrow \emptyset$ 
for all  $I \in \mathcal{V}$  do
     $\mathcal{K} \leftarrow predict(\mathcal{K})$ 
     $E \leftarrow eyeDetect(I)$ 
     $C \leftarrow dist(E, position(\mathcal{K}))$ 
     $m_E, m_{\mathcal{K}} \leftarrow JonkerVolgenant(C)$ 
    for all  $i \in [1, S]$  do
         $\mathcal{K}[m_{\mathcal{K}}[i]] \leftarrow update(\mathcal{K}[m_{\mathcal{K}}[i]], E[m_E[i]])$ 
    end for
     $i_{unmatched\_trackers} \leftarrow \{i \in \mathbb{N} | i \leq M, i \notin m_{\mathcal{K}}\}$ 
    for all  $i \in i_{unmatched\_trackers}$  do
         $\mathcal{K}[i] \leftarrow incrUnmatched(\mathcal{K}[i])$ 
        if  $unmatched(\mathcal{K}[i]) > k$  then
             $\mathcal{K}_{log} \leftarrow \mathcal{K}_{log} \cup \mathcal{K}[i]$ 
             $\mathcal{K} \leftarrow \mathcal{K} \setminus \mathcal{K}[i]$ 
        end if
    end for
     $i_{unmatched\_detections} \leftarrow \{i \in \mathbb{N} | i \leq N, i \notin m_E\}$ 
    for all  $i \in i_{unmatched\_detections}$  do
         $\mathcal{K} \leftarrow newTracker(E[i]) \cup \mathcal{K}$ 
    end for
end for
Return  $\mathcal{K} \cup \mathcal{K}_{log}$ 

```

---

### 2.3. Module 3: semi-automated labeling

The algorithm described in Subsection 2.2 can track and differentiate salmon individuals as long as they reside in the camera field of view. However, it cannot re-identify salmon across sessions since it is merely based on the continuity of motion, and does not utilize any appearance features. To re-identify individual salmon over a longer period, a salmon operculum and anterior trunk (OAT) dataset was created, consisting of salmon images rectified to show the fish in a similar pose, each associated with an individual label describing the fish individual (Fig. 3). The raw OAT images were created by using the four most posterior keypoints on each salmon, received from the Keypoint RCNN model described in Subsection 2.1, to calculate and apply a homogeneous transformation (homography) (see e.g. (Hartley and Zisserman, 2004)) that mapped all fish onto a fixed configuration. We note that for each fish we consider both sides of the salmon. Subsequently, each transformed salmon image was cropped to contain only the posterior head and anterior trunk of the fish, outlined by the dorsal fin, pectoral fin and eye keypoints. When annotating the fish images with 14 class labels, algorithm 1 was first employed to extract an OAT image of the same fish from each frame in a trajectory, all linked to a particular tracker identifier. This reduced the

A dataset of 1014 OAT images from each fish class was created, with each class containing images from videos recorded on different days. Additionally, a test dataset was created from novel trajectories containing 105 images per class. The training dataset was inflated thirtyfold by an Albumentations (Buslaev et al., 2020) pipeline consisting of random tone curve perturbations, square dropouts, blurring, grid distortion and brightness and contrast adjustment.

### 2.4. Module 3: salmon re-identification

Having constructed a salmon OAT dataset for supervised salmon re-identification (Subsection 2.3), a neural network classification architecture can be trained to recognize individual salmon. In this work, a ResNet 101 (He et al., 2016) PyTorch (Paszke et al., 2019) model with a pretrained backbone and a randomly initialized 14 class network head (final network layer) was trained on the OAT dataset. The model was trained for 8 epochs using a learning rate of 0.005, a momentum of 0.9 and a weight decay of 0.0005 after four epochs. This training took around 7 h on hardware consisting of 2 T4 GPUs and 1 Intel Xeon CPU. The accuracy of the model was evaluated on the test data of the OAT dataset. Two additional networks, AlexNet (Krizhevsky et al., 2012) and Efficientnet (Tan and Le, 2019), as well as two classical methods, scale-invariant geometric relationship between anatomical landmarks and

**Table 1**

An example of two respiration frequency ranking lists, and the corresponding respiration frequency ranking change between them. E-M is an abbreviation for Ebed-Melech.

Respiration frequency ranking list 051022, camera 7, clip 16	E-M	Hosea	Daniel	Caiphas	Aaron	Gehasi	Baasha
Respiration frequency ranking list 051022, camera 7, clip 17	E-M	Daniel	Hosea	Caiphas	Aaron	Gehasi	Baasha
Respiration frequency ranking change	0	1	1	0	0	0	0

melanin spot detection followed by direct dot set comparisons, were tried out, but showed lower performance than the selected model.

### 2.5. Module 4: mouth frequency analysis

The trackers (found by [algorithm 1](#)) contain information about the fish id (class) and the salmon jaw gape for each frame over a short video recording. This information can be interpreted as a set of time series, where the time stamp corresponds with the frame number, and the associated value is a tuple containing the fish class (one of the 14 possible classes) and the Euclidean distance between the upper and lower jaw. Other jaw gape descriptors were tested, such as Euclidean jaw distance after homogeneous transformation ([Subsection 2.3](#)) and mouth opening angle, however the raw Euclidean distance was found to provide the best input for the downstream modules. When discussed separately, the class data of a time series will be referred to as a class list, and the jaw gape data as a jaw gape time series. As the time series received from [algorithm 1](#) were of unequal length, and often included outliers, the following two preprocessing steps were performed:

1. Removal of clear outliers in the time series. This was performed by removing keypoints that had a clearly erroneous relative placement. The function that performed this pruning was called `remove_err_keypoints() : K[i] → K[i]`.
2. Extracting fixed length (L) time series from the raw data by using a fixed size sliding window. This preprocessing step helped to increase redundancy in the salmon respiration frequency estimation method and reduce fluctuations in the jaw gape time series.

Both a respiration frequency and a single class value (providing the fish ID) were extracted from each of these L-length time series. The fish class was determined by using the most common class that appeared in the L-length class list (*mode:  $\mathbb{R}^L \rightarrow \mathbb{R}^1$* ), while the respiration frequency was found by extracting the frequency of a sine wave fitted to the jaw gape time series. The employed sine wave function ([Eq. \(1\)](#)), contains five parameters that can be adjusted in order to capture the dynamics of the jaw gape time series: amplitude (A), frequency (f), phase ( $\phi$ ), offset (o) and offset inclination (a). Some traditional methods for signal frequency estimation were evaluated as well, such as autocorrelation and Fourier transformations, however they were not used as they could not deal with the complex dynamics of the jaw gape time series.

The fitting of the sine function ([Eq. \(1\)](#)) to jaw gape measurements was solved by exploiting the Levenberg-Marquardt method, which is a gradient descent approach that converges to a local minima ([Lourakis, 2005](#)). To ensure removal of the remaining outliers, the Levenberg-Marquardt fitting was used inside a random sample consensus (RANSAC) loop ([Raguram et al., 2013](#)), which randomly samples a minimal set of data points from the jaw gape time series, fits a sine function to these by the Levenberg-Marquardt method, counts the number of time series points with distance below a threshold to the estimated model and returns the model with fewest outliers after having evaluated a certain

number of point samples. As the Levenberg-Marquardt method requires initial parameter guesses, the RANSAC algorithm was repeated for a set of initial frequency parameters (*freqVal*), to ensure frequency estimates free of initialization bias. The *freqVal* set contained equally spaced values covering expected salmon frequencies. The rest of the initial parameters ( $\mathbf{p}_0$ ) were randomly sampled around expected values calculated from jaw gape time series statistics such as the average jaw gape value ( $O_0$ ), the inclination of a straight line fitted to the jaw gape data ( $a_0$ ) or the weighted (multiplied by  $\sqrt{2}$  ([Li et al., 2019](#))) standard deviation between this line and the jaw gape data ( $A_0$ ). Furthermore, the RANSAC inlier threshold was set to this standard deviation without the correction weight, and  $\phi_0$  was randomly sampled in the range  $[0, 2\pi]$ .

$$f(x, \mathbf{p}) = A \cdot \sin(2\pi \cdot f \cdot x + \phi) + o + a \cdot x \quad (1)$$

The reliability of the respiration frequency estimate of each time series was assessed by the function `acceptable_time_series() : TS → {0, 1}`. Estimates were discarded if the fish classification was ambiguous, if there were any data integrity issues, or if there were unexpected anomalies in the measurement process. This assessment was based on the following four criteria:

1. The predominant salmon class within a time series must be statistically significant, i.e., one salmon class must clearly dominate the rest.
2. The number of classification estimates remaining in a class list, after filtering out uncertain classifications, must be sufficient. Certainty scores derived from the ResNet 101 model ([Subsection 2.4](#)) were utilized for the filtering process.
3. Time series must not exhibit an excessive number of erroneous keypoint predictions. This criterion addresses the issue of data integrity, ensuring that only reliable data points contribute to the analysis.
4. The sine model with the highest inlier count, obtained from the RANSAC loop, must not have a low number of inliers, indicating poor model fit.

The complete respiration frequency extraction method is explained in [algorithm 2](#), where  $\mathcal{F} = \{(c, f) : c \in \mathbb{N}, f \in \mathbb{R}^+\}$  is a set of tuples with class and frequency estimates,  $\mathcal{G}$  is a set of L-length time series (containing jaw pose and individual class information for each frame),  $\mathbf{p} \in \mathbb{R}^5$  contains sine wave parameters,  $\sigma_f^2 \in \mathbb{R}^1$  is the frequency variance of a Levenberg-Marquardt fit,  $\mathbf{IP}$  contains RANSAC inlier points,  $\mathbf{p}_f \in \mathbb{R}^1$  is a sine wave frequency parameter,  $\mathbf{TS}_g \in \mathbb{R}^{2 \times L}$  is a jaw pose time series and  $\mathbf{TS}_c \in \mathbb{R}^L$  is a class list.

After finding the fish class and respiration frequency for each L-length time series (the set  $\mathcal{F}$  of [algorithm 2](#)), the frequency data from video recordings are aggregated into a singular frequency estimate for each fish. This aggregation is achieved by calculating the median of all available respiration frequency measurements for each individual fish.

**Algorithm 2.** Salmon respiration frequency estimation from time series

---

```

Require:  $\mathcal{K}$ , freq_val
 $\mathcal{F} \leftarrow \emptyset$ 
for all  $\mathcal{K}[i] \in \mathcal{K}$  do
     $\mathcal{K}[i] \leftarrow \text{remove\_err\_keypoints}(\mathcal{K}[i])$ 
     $\mathcal{G} \leftarrow \text{moving\_window}(\mathcal{K}[i])$ 
    for all  $\text{TS} \in \mathcal{G}$  do
         $IP_{\max} \leftarrow 0$ 
        for all  $f_0 \in \text{freqVal}$  do
             $\mathbf{p}_0 \leftarrow [A_0, f_0, \phi_0, O_0, a_0]$ 
             $\mathbf{B}, \mathbf{IP}, \sigma_f^2 = \text{ransac}(\text{TS}_g, \mathbf{p}_0)$ 
            if  $\text{dim}(\mathbf{IP}) > IP_{\max}$  and  $\sigma_f^2 < \text{maxFreqStd}$  then
                 $\mathbf{p}^* \leftarrow \mathbf{p}$ 
                 $IP_{\max} \leftarrow \text{dim}(\mathbf{IP})$ 
            end if
        end for
         $c \leftarrow \text{mode}(\text{TS}_c)$ 
        if  $\text{acceptable\_time\_series}(\text{TS})$  then
             $\mathcal{F} \leftarrow \mathcal{F} \cup \{(c, \mathbf{p}_f^*)\}$ 
        end if
    end for
end for

```

---

The underlying hypothesis that jaw gape oscillations reflect the respiration frequency of salmon is based on three assumptions:

1. **Descriptiveness of jaw gape oscillations:** The frequency of the jaw gape pattern correlates largely with the salmon ventilation frequency.
2. **Sinusoidal motion of jaw gape:** The amplitude variation between the upper and lower jaws follows an almost ideal sine wave pattern, maintaining a consistent frequency throughout the observation period.
3. **Jaw gape frequency band separability:** The motion dynamics of the entire salmon contain no major frequency components in the same range as the salmon ventilation frequency.

If these assumptions hold, the 1–2 Hz oscillations observed in the jaw gape time series (Section 3) approximate the salmon respiration frequency, while the offsets, amplitude variations and inclinations of the jaw gape time series reflect the motion and position of the entire salmon.

## 2.6. Experiment

The capability of the proposed method (SaBRE) is explored and assessed based on data obtained during a biological driven lab-scale salmon stress experiment. The full video-data-set was gathered from 9 water tanks (1x1x1 meter) with a cluster of three tanks in one of three different temperature ranges. The tanks contained saltwater, and were oxygenated by air stones and water flow into the tanks. Each tank contained 7 fish which were fed regularly by automatic feeders. For demonstrating the developed technology we showcase the developed method on data from one of these tanks. In this tank, the weight of the fish population was  $0.74 \pm 0.06$  kg and the fork length of the fish popu-

lation was  $40.43 \pm 1.15$  cm. The protocols for handling and care of salmon followed in this study were reviewed and approved by the Norwegian Food Safety Authority (NFSA), approval number FOTS ID 29871 and were conducted following the ethical guidelines from the NFSA.

Following an initial acclimatization phase, the fish were maintained in a tank for a duration of approximately 14 days. During the first phase, which lasted for 6 days, the water temperature was gradually increased from  $12.0^\circ\text{C}$  to  $15.9^\circ\text{C}$ . Subsequently, on the 7th day, a respiratory experiment was conducted. During this experiment, the oxygen supply to the tank was restricted by stopping the inflow of water and by removing the air stone, which is normally submerged to facilitate oxygenation. Consequently, due to the breathing of the fish, the dissolved oxygen ( $\text{dO}_2$ ) content within the tank decreased, as measured by a *miniDO<sub>2</sub>TLogger* (PME, 2021). Upon reaching an oxygen concentration of 45 %, both water and air were reintroduced into the tank. This action incrementally restored the  $\text{dO}_2$  levels to their initial range, approximately 94 %, as illustrated in Fig. 7. During the next 6 days, the temperature was held relatively constant and was at  $17.7^\circ\text{C}$  on the final day of the experiment when a new down-breathing was performed and where the  $\text{dO}_2$  content reached a minimum of 49 %.

During the down-breathings, the tank was filmed with the help of two submerged GoPro cameras in fixed, horizontal orientation 21 cm over two bottom corners of the tank, operating at 60 frames per second. The tank was illuminated by a LED spotlight placed above the tank. Roughly once every hour, the batteries and memory cards of the cameras were switched, introducing disturbances in the tank and recording breaks in the video data. We call the period between two battery changes one *session*. A *session* is approximately one hour long and is split into consecutive clips of 707 s (42,420 frames). The length of the final clip of each session varies according to the time of battery change.

In this way, we obtained four video *recording sets* containing data from the two cameras employed during the two down-breathings. Each of these *recording sets* was processed separately.

For each clip in a *recording set*, frames 5000 to 7000 were passed through SaBRE, providing information on the breathing frequency development over time for each individual. In addition to this fine-grained *full frequency evolution*, a *session frequency evolution* was calculated for each fish by pooling the  $\mathcal{F}$  set (see [algorithm 2](#)) of all clips in a session, and calculating the median respiration frequency for each fish over this extended set.

The first clip of each session was discarded when calculating the *session frequency evolution* due to an observed respiration frequency increase of the fish after camera insertion. In order to ensure robust results for both the full and session frequency evolutions, a fish was considered absent at a given time instance if the number of respiration frequency estimates over a clip or a session (the size of the  $\mathcal{F}$  set) were under a threshold. Furthermore, by averaging over the fish individuals at each time interval (clip or session), both an *average full frequency evolution* and an *average session frequency evolution* were created.

To evaluate the effects of camera insertion, the change in respiration frequency between the clip of camera insertion and the temporally adjacent clips was calculated. The nature of the disturbances differ both in magnitude and length, however we hypothesize that the fish has a significant disturbance response at the SaBRE evaluation of the first clip of each *session* (up to 7000 frames after the disturbance has stopped), and that the disturbance response becomes negligible after one video clip. Eq. (2) displays the formula for calculating the *respiration frequency*

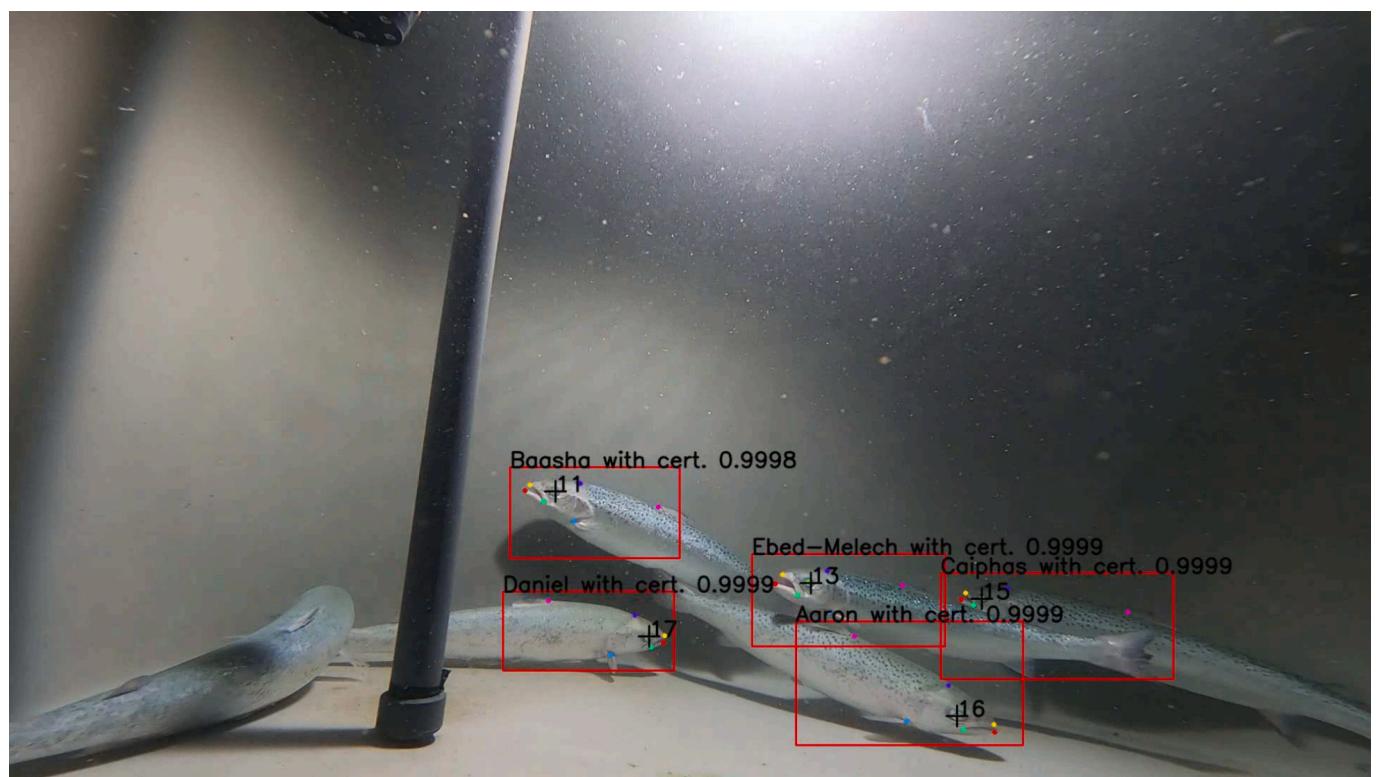
*increase after disturbance* at a disturbance time  $t$ . Here,  $\mathbf{f}$  is an *average full frequency evolution*,  $t$  is a clip index, and *session\_start* is a set containing the indices of all clips beginning a session.

$$\mathbf{f}_{inc,t} = \begin{cases} \frac{(\mathbf{f}_t - \mathbf{f}_{t-1}) + (\mathbf{f}_t - \mathbf{f}_{t+1})}{2} & \text{if } t > 1, t \in \text{session\_start} \\ \mathbf{f}_t - \mathbf{f}_{t+1} & \text{if } t = 1 \end{cases} \quad (2)$$

Finally, the consistency of respiration frequency rankings of individual salmon was evaluated by the metric *respiration frequency ranking change*. This metric evaluates to what extent a salmon group has the same respiration frequency ranking at two different times and is calculated the following way ([Table 1](#) for an example).

1. Extract the respiration frequency ranking list at two different time points. A respiration frequency ranking list contains the name of each salmon in a group, ordered according to their respiration frequency value.
2. Calculate how many places each salmon rises or falls between the two lists.

The *respiration frequency ranking change* metric was calculated for the following respiration frequency lists, which also include a baseline based on simulated randomly ordered lists:

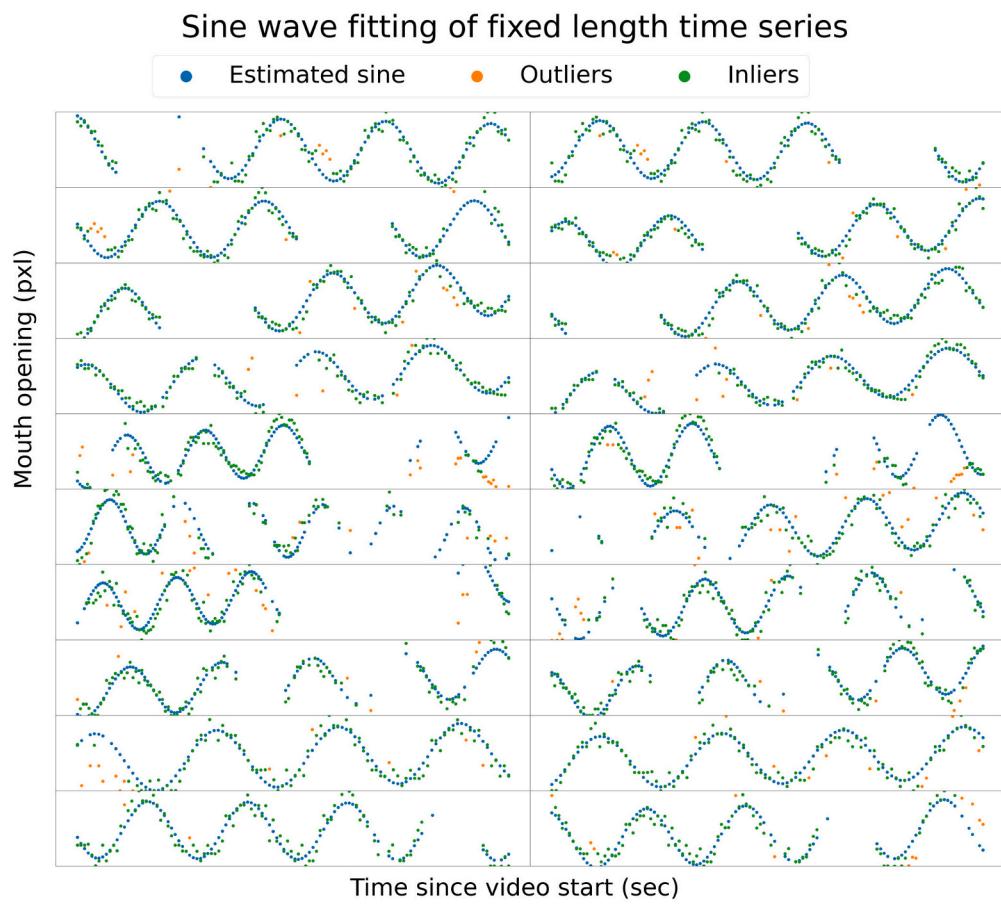


**Fig. 4.** A video from the salmon down-breathing experiment annotated with the results of modules 1, 2 and 3 of our salmon breathing rate estimation (SaBRE) pipeline. The black object at the top of the frame is the  $dO_2$  sensor. Stating that the animated video can be viewed in the supplementary material.

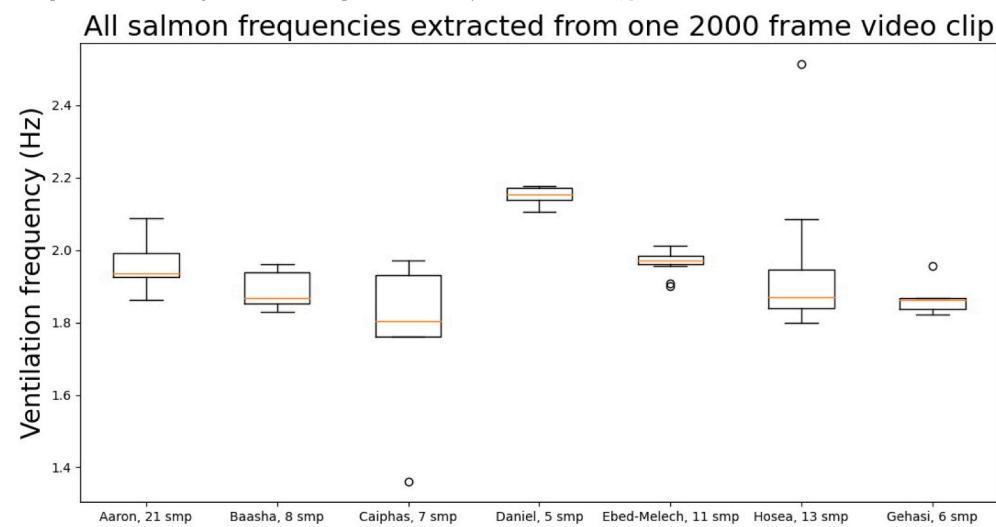
**Table 2**

Occurrences of different tracker behaviors. The clip name follows the following convention: “date (ddmmyy), tank number, camera number, session number, clip number (in given session)”.

Clip	Correct tracks	Premature termination	Trackers switching fish
051022, t9, c7, 2 t, s3	22	0	0
051022, t9, c8, 2 t, s3	24	1	5
280922, t9, c8, 2 t, s3	14	2	2
280922, t9, c9, 2 t, s3	16	3	3
Sum	76	6	10

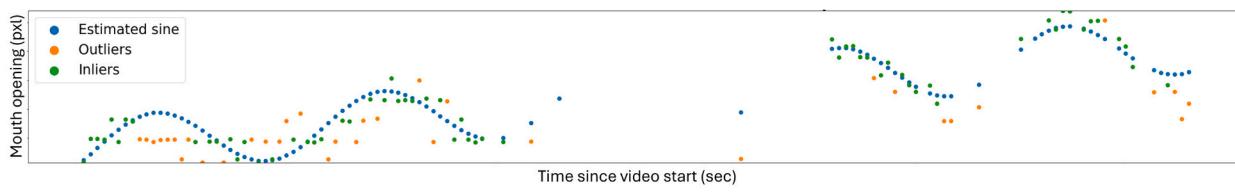
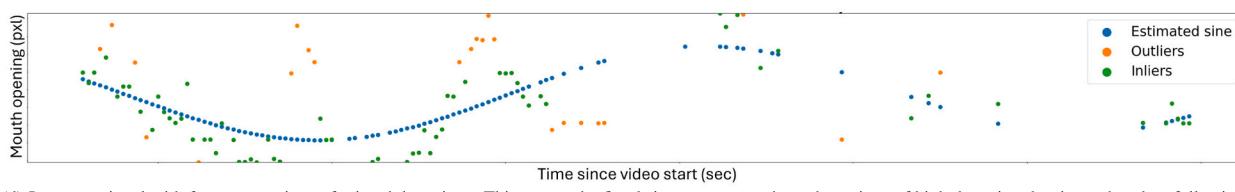
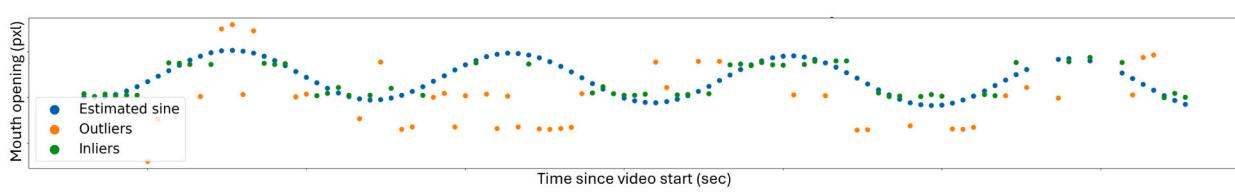
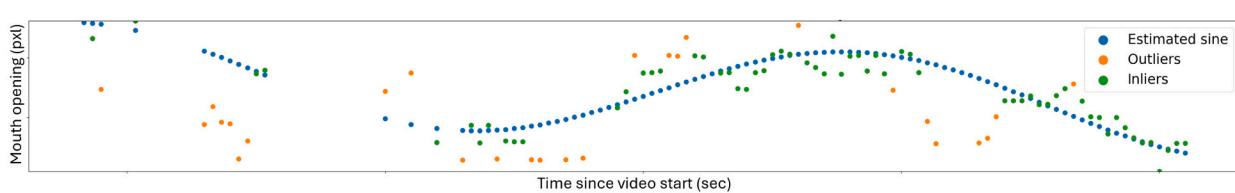
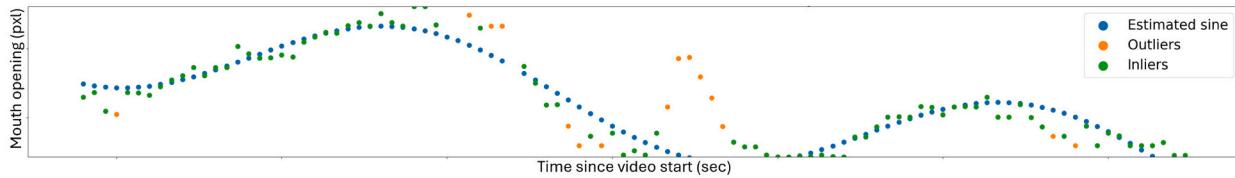


(a) A representative sample of jaw gape time series fitted to sine waves. All axes have different scales and different offsets, hence only the frequencies should be compared across the subplots. Each subplot contains 100 jaw gape points, and the time span displayed in each plot differs according to the number of points removed by the `remove_err_keypoints()` function.



(b) All respiration frequencies extracted from a representative 2000 frame video clip, ordered in a box plot according to the assigned fish class. The number of samples per fish can be read as the smp value, and the orange lines are the median fish respiration frequencies (SaBRE output).

**Fig. 5.** Results from module 4.



**Fig. 6.** Examples of the cases where SaBRE was not able to correctly estimate salmon ventilation frequency from a jaw gape time series. The green and orange dots are the inliers and outliers estimated by the RANSAC method, respectively. The blue dots are samples of the estimated sine wave model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 3**

Number of correct and erroneous salmon respiration frequency estimates, evaluated by comparing jaw gape time series, sine wave model fits and annotated videos. The clip name follows the following convention: “date (ddmmyy), tank number, camera number, session number, clip number (in given session)”.

Clip	Correct	Irregular respiration	Body motion	Keypoint noise	Detection noise	Tracker switching fish
051022, t9, c7, 2 t, s3	63	3	3	2	0	0
051022, t9, c8, 2 t, s3	29	2	0	0	0	2
280922, t9, c8, 2 t, s3	125	2	1	7	1	1
280922, t9, c9, 2 t, s3	98	1	1	4	0	1
Sum	315	8	5	13	1	4

- Session respiration frequency ranking lists.** These ranking lists were extracted from the median session respiration frequencies (data points in the *session frequency evolutions*). The ranking changes were calculated between consecutive sessions.
- Down-breathing respiration frequency ranking lists.** These ranking lists were extracted from the average respiration frequency of all session respiration frequencies in a down-breathing recording. The ranking changes were calculated between the two down-breathings.
- Randomly ordered ranking lists (baseline).** These rankings lists were extracted as a random list of salmon names. The ranking changes were calculated by simulating 84 random ranking pairs. This simulation was repeated 1000 times, resulting in both a mean and a standard deviation of the average ranking change of random ranking lists.

### 3. Results

#### 3.1. Salmon breathing rate estimation (SaBRE)

Quantitative assessment of the explained Keypoint RCNN model is performed in He et al. (2016) and agrees with the effectiveness we observe for the Keypoint RCNN model applied on our data.

All analyzed video clips were annotated with bounding boxes, keypoints, trackers, and salmon identifiers to facilitate assessments of the SaBRE modules. An example video is shown in Fig. 4. To validate the tracker performance, we review a selection of these videos, and observe three distinct tracker behaviors:

- The tracker spawns correctly when a salmon eye is detected, and terminates correctly when the salmon eye has been occluded for the specified number of frames. A tracker is considered spawned when it has been linked to ten detections of the same salmon.
- The tracker terminates prematurely, i.e. before the eye has been occluded for the specified number of frames. A tracker is considered prematurely terminated if it does not match with a detection (of the salmon it follows) that happens less than ten frames after the last match of the tracker. A variant of this behavior is when a single salmon causes two salmon detections, resulting in the spawning of a second tracker very close to the original one (see tracker 18 in the video of Fig. 4). If the newly spawned tracker ends up closer to the true salmon than the old one, the old tracker will terminate prematurely.
- The tracker switches salmon. This happens when two salmon eyes are close together, and the tracker is matched with the detection of a new salmon. A tracker is considered to have switched salmon if it has been linked to two separate salmon for ten frames each.

To quantify how often each of these cases occurs, we sample four annotated video clips, one from each *recording set*, and count the occurrences of the different tracker behaviors (Table 2). Premature termination is not very problematic, as this only causes a slight reduction in the amount of usable data for later modules. Trackers switching

fish, however, could lead to erroneous matching between salmon identity and respiration frequency. For the random sample we analyzed, around 11% of trackers belong to this problematic error group (Table 2).

The salmon re-identification module was evaluated by calculating the performance of our re-identification models on the test data consisting of novel salmon trajectories. The evaluation of the best model, ResNet 101, resulted in a high accuracy of 99.51 %.

The error cases of module 3 (salmon re-identification) had a tendency to be between visually similar classes (such as Gehasi and Daniel), and for the most distorted images (due to inaccurate keypoint detections and subsequent homogeneous transformation, see subsections 2.1 and 2.3).

To validate the performance of module 4, plots displaying estimated sine models and jaw pose time series data (Fig. 5a) were compared with annotated videos. The same four video clips as in the tracker assessment were evaluated. We observed five different situations where SaBRE was unable to correctly estimate the salmon ventilation frequency:

- Irregular respiration** (Fig. 6a): Major, short term variability in the jaw gape frequency of the salmon. Mainly due to salmon snapping. Violating assumption 2 (sinusoidal motion of jaw gape) of Subsection 2.5.
- Body motion** (Fig. 6b): Excessive salmon motion, causing major frequency components in the jaw gape time series that do not correspond to ventilation frequency. Violating assumption 3 (jaw gape frequency band separability) of Subsection 2.5.
- Keypoint noise** (Fig. 6c): Keypoint localization errors dominating the ventilation frequency information in the jaw gape time series. Caused by a weakness in module 1. Some of the jaw gape time series generated by Daniel in the video of Fig. 4 gave erroneous salmon respiration frequency estimates due to keypoint localization noise.
- Detection noise** (Fig. 6d): Frequent frame sections of missed detections. Caused by a weakness in module 1.
- Trackers switching fish** (Fig. 6e): Trackers switching fish, causing the respiration frequency originating the jaw gape time series to be nonunique. Caused by a weakness in module 2.

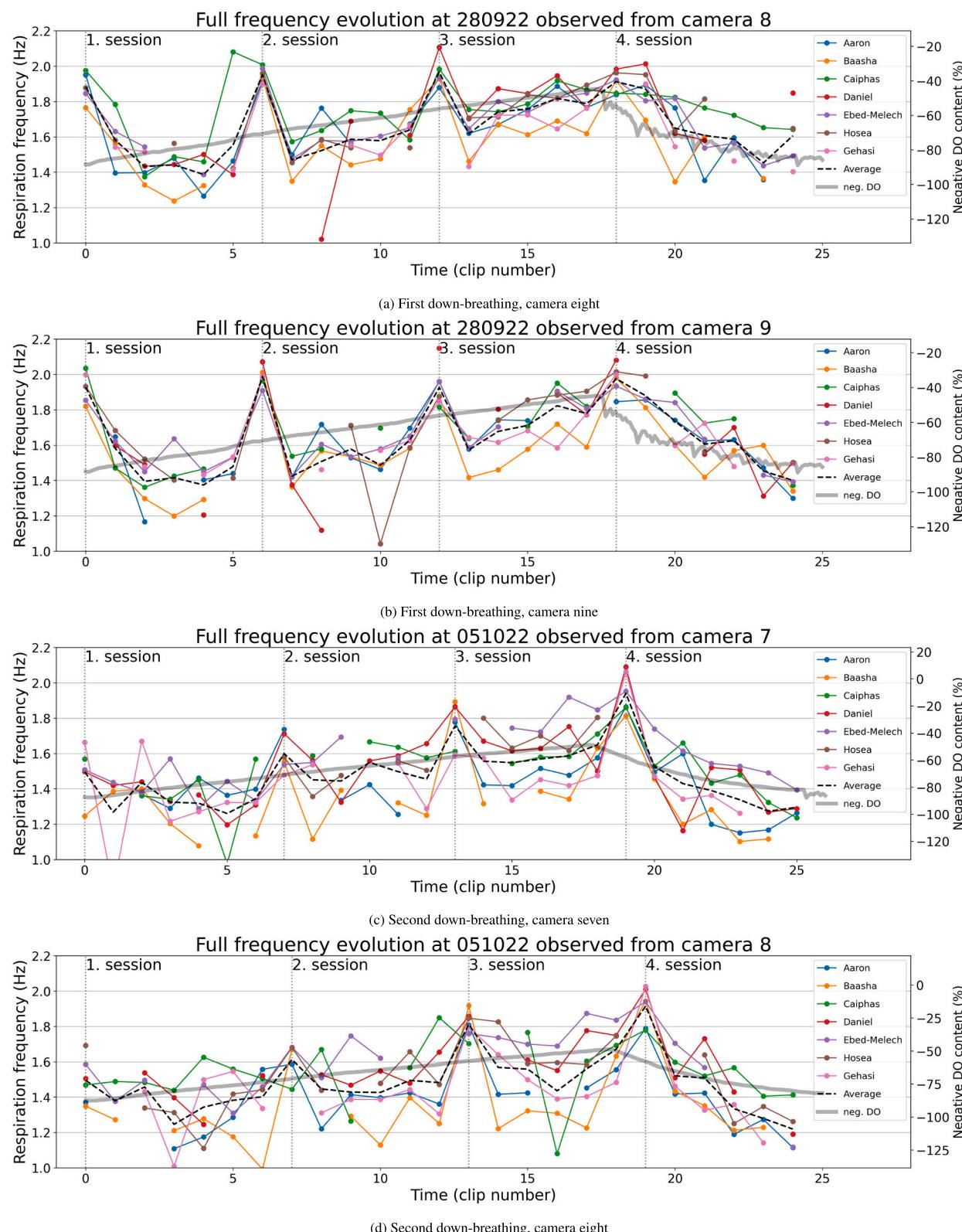
Around 91 % of the salmon ventilation frequency estimates were correct on the evaluated sample (Table 3). Of the 31 errors we observed, 13 came from assumption violations (irregular respiration and body motion), while 18 came from method shortcomings (keypoint noise, detection noise and trackers switching fish).

In Fig. 5b, a box plot displaying the extracted breathing frequencies for each fish individual over a short video recording (clip) is displayed. The median respiration frequency of each fish, shown as an orange line, is the final SaBRE respiration frequency output.

#### 3.2. Frequency evolutions

Full frequency evolutions obtained from all four *recording sets* (two down-breathings filmed with two cameras) are displayed together with the negated dO<sub>2</sub> content at the times of recording in Fig. 7.

The dO<sub>2</sub> content and the respiration frequency were plotted in the



**Fig. 7.** Full frequency evolutions for both cameras during both down-breathings. The dashed, vertical gray lines are added at the disturbance clips, the colored trajectories display the respiration frequencies of individual fish, the dashed, black trajectories display the average full frequency evolutions, and the partially transparent, thick gray trajectories display the negated  $\text{DO}_2$  content.

**Table 4**

Pearson correlation coefficients ( $r$ ), and the likelihood of them being different from zero ( $p$ ), between all *average full frequency evolutions* (without disturbance clips) and the  $\text{dO}_2$  content at the times of SaBRE evaluation.

Date and camera	$r$	$p$
28.09.22, camera 8	-0.67	8.6e-04
28.09.22, camera 9	-0.61	3.4e-03
05.10.22, camera 7	-0.90	7.3e-09
05.10.22, camera 8	-0.78	3.1e-05

same value range by matching the initial  $\text{dO}_2$  content with the first *average session frequency evolution* value, and the maximum negated  $\text{dO}_2$  content with the highest non-disturbance value of the *average full frequency evolution*.

As the cameras employed in this experiment did not provide accurate time stamps for the video recordings, the oxygen data and the respiration data must be aligned in the temporal dimension. To align the two data sources, we matched both of them to the start of the down-breathing experiment, and assumed each clip to be 12 min long. This might induce an offset (estimated to be between  $\pm 5$  min and  $\pm 10$  min), mainly due to the temporal drift incurred during the manual battery changes of the cameras in the tank, and is the most likely explanation for the discrepancy between the time of maximum respiration frequency and minimum  $\text{dO}_2$  content.

One way to evaluate the respiration frequency estimates is to compare the frequency predictions for a particular fish and clip number between cameras. The differences observed in such a comparison can have two causes: a SaBRE error, or variability in the physical salmon respiration frequency over the time span of a video clip (about 33 s). As methods for removing frequency estimates not corresponding to physical salmon jaw gape frequencies are already implemented in SaBRE (Subsection 2.5), no further automatic methods are directly available for determining which of these two sources a potentially observed discrepancy is caused by. What we can say, however, is that some of the respiration frequency estimates that initially appear as outliers are, in fact, approximately equal in both cameras, like Daniel in the eighth clip during the first down-breathing (Fig. 7). This indicates that salmon occasionally have significant respiration frequency changes during a video clip, which would, together with different view directions of the two cameras, explain variability in frequency predictions even if the SaBRE algorithm would have had no errors.

It seems reasonable to assume that a combination of the SaBRE errors addressed in Subsection 3.1 and respiration frequency changes during a video clip are responsible for the observed respiration frequency discrepancies in the two simultaneously filming cameras discussed in the paragraph above.

Significant and consistent correlations were detected between  $\text{dO}_2$  and respiration frequency, with Pearson correlation coefficients ranging from -0.61 to -0.90 (Table 4). As such, we conclude that our data indicates a negative correlation between  $\text{dO}_2$  content and respiration frequency. This corresponds well with the trends observed in Fig. 7.

### 3.3. Salmon rankings

Table 5 presents the respiration frequency ranking changes calculated in this paper. The numbers in the first row specify the baseline ranking changes (r.c.) computed between many pairs of randomly

created respiration frequency ranking lists. The numbers in the subsequent rows show the rate of ranking changes observed in the videos in the tank. The “avg. r.c.” column displays the average ranking change, the “smp.” column displays the number of ranking changes used to calculate the ranking change rates, and the “exp. num.” column displays the number of times the experiments were repeated. Since only one salmon tank is considered in this paper, the experimental rankings (session and down-breathing) have an experiment repetition number of one.

Both the session ranking lists and the down-breathing ranking lists have an average ranking change below that of the randomly ordered lists, which means that the data of our experiment suggest that two experimental respiration frequency ranking lists tend to be relatively similar (i.e., the salmon that respire slowest at a given SaBRE gauging tend to be among the salmon displaying the lowest respiration frequencies also at different SaBRE evaluations).

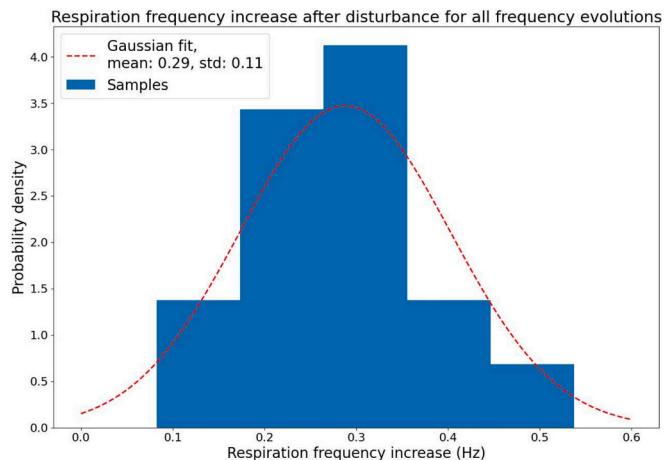
The sizeable average ranking change discrepancy between the experimental and random rankings, together with the high sample number and the low uncertainty of the average ranking change of the randomly ordered lists, allows a reasonable expectation that the experimental results observed above are generalizable to other fish groups and settings. From the available data at our disposal, we can infer that the ranking of individual salmon is mostly consistent over both the short (session) and long (down-breathing) term.

### 3.4. Respiration frequency increases after disturbance

Respiration frequency was observed to consistently and significantly increase in response to camera insertion (Fig. 8).

### 3.5. Resource requirements

The processing time of generating jaw gape time series from one 2000 frame image sequence was 90 min, while the frequency fitting (algorithm 2) took 2 min. This yielded a total processing time of 37 h for one down-breathing when using an Intel i7 processor.



**Fig. 8.** Histogram showing the change in respiration frequency at clips with camera insertion.

**Table 5**

Rate of ranking changes for all salmon ranking lists (truncated after 3 digits). For the average ranking change of the random salmon ranking, the uncertainty is presented as twice the standard deviation ( $\pm 2\sigma$ ).

	0 r.c.	1 r.c.	2 r.c.	3 r.c.	4 r.c.	5 r.c.	6 r.c.	Avg. r.c.	Smp.	Exp. num.
Randomly ordered ranking lists	0.143	0.245	0.204	0.163	0.123	0.082	0.041	2.286 $\pm$ 0.134	84	1000
Session respiration frequency ranking lists	0.167	0.381	0.179	0.119	0.119	0.023	0.012	1.762	84	1
Down-breathing respiration frequency ranking lists	0.214	0.357	0.214	0.071	0.143	0.000	0.000	1.571	28	1

#### 4. Discussion and future work

This paper introduces SaBRE, an algorithm capable of monitoring the respiration frequency of individual salmon from video data. By employing our method to analyze data from a salmon down-breathing experiment, it has been demonstrated that the algorithm estimates respiration frequencies that correlate with changes in the environment. These include disturbances, correlating with a short-term respiratory increase, and reduced dO<sub>2</sub> content, correlating with increased respiratory frequency. Additionally, the respiration frequency ranking between fish in the same tank appears to be relatively constant both over the short and long term.

Both the presence of a negative correlation between dO<sub>2</sub> content and respiration frequency in salmon (Remen, 2012) and a positive correlation between disturbances and respiration frequency in fish (Davis and Schreck, 1997) are well established, although some details are not fully understood. The consistency of fish respiration frequency ranking is not as well researched, and there are several possible explanations for the fish behavior observed in our experiment, including different coping styles (Barreto and Volpato, 2011) and different weights (Magid and Babiker, 1975) of the individual salmon. By employing SaBRE to analyze datasets from larger experiments, details regarding the relationships indicated by our data can be properly established and examined, further progressing the understanding of salmon physiology.

We have demonstrated that all SaBRE modules are highly accurate. Our re-identification module achieves an accuracy of 99.51 %, our tracker module shows no fish switches in 89 % of cases, and our mouth frequency analysis method is able to correctly estimate jaw gape frequencies for 91 % of the evaluated jaw gape time series. The performance of our mouth frequency estimation module (Subsection 3.1) demonstrates that our assumptions about sinusoidal motion of jaw gape and jaw gape frequency band separability (Subsection 2.5) are reasonable, with a combined violation rate of under 4 %. As our method considers the median respiration frequency across video clips to calculate individual fish respiration frequency, it is unlikely that the few respiration frequency errors caused by assumption violations have any significant impact on the final respiration frequency estimates.

The biological experiment was well constructed, and allowed examination of three interesting biological effects. Even more powerful comparisons between different down-breathing recordings could have been made if timing information had been available in the video clips. This would be an easy addition in future experiments, by starting each video by filming a common clock.

In the future, we aim to employ SaBRE in large-scale industrial and scientific settings. The most prominent challenge to such a deployment is the problem of re-identifying hundreds of thousands of salmon, which would require a considerable improvement of current methods of salmon re-identification. SaBRE could be deployed without the salmon re-identification branch, which would allow continuous monitoring of the group salmon respiration frequency, but would remove the possibility of assessing long-term respiration frequency evolutions of individual fish. Another limiting factor of SaBRE is that our algorithm assumes respiration by buccal pumping, which limits the application of the methodology to salmon swimming at low speeds.

In conclusion, this paper demonstrates the feasibility of computer vision-based respiration monitoring of individual salmon, and shows that the data extracted by such a framework contains exciting patterns that can be used as indicators of important farm states. We expect SaBRE to have the potential to contribute to a significant improvement to aquaculture operations by providing valuable insight into the current fish conditions. This will lead to more effective, humane and sustainable farm sites, bringing salmon farming closer to its full potential.

#### 5. Declaration of use of generative artificial intelligence (AI)

Statement: during the preparation of this work the author(s) used

Grammarly in order to locate and improve potentially problematic grammar and wording. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aquaculture.2024.741535>.

#### CRediT authorship contribution statement

**Espen Berntzen Høgstedt:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

**Christian Schellewald:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

**Rudolf Mester:** Writing – review & editing, Supervision. **Annette Stahl:** Writing – review & editing, Supervision.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Christian Schellewald reports financial support was provided by Norwegian Seafood Research Fund. Christian Schellewald reports financial support was provided by Research Council of Norway. Espen Berntzen Høgstedt reports financial support was provided by Research Council of Norway. Espen Berntzen Høgstedt reports financial support was provided by Norwegian Seafood Research Fund. Annette Stahl reports financial support was provided by Research Council of Norway. Rudolf Mester reports financial support was provided by Research Council of Norway. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

The data that has been used is confidential.

#### Acknowledgment

This work was supported by the Norwegian Seafood Research Fund (FHF) under project number 901736, titled 'Kunnskapsgrunnlag for biologisk relevante velferdsindikatorer for laks i akvakultur (BIORELEVANS)',; and by the Research Council of Norway (RCN) under project number 344022, titled 'Computer Vision and Artificial Intelligence based Salmon Identification and automated long-term welfare assessment in aquaculture net-pens (cAIGe)'.

#### References

- Alver, M.O., Føre, M., Alfredsen, J.A., 2022. Predicting oxygen levels in Atlantic salmon (*Salmo Salar*) sea cages. Aquaculture 548, 737720. URL: <https://doi.org/10.1016/j.aquaculture.2021.737720>.
- Aota, S., Holmgren, K.D., Gallaugher, P., Randall, D.J., 1990. A possible role for Catecholamines in the Ventilatory responses associated with internal Aciosis or external hypoxia in rainbow trout *Oncorhynchus Mykiss*. J. Exp. Biol. 151, 57–70. URL: <https://doi.org/10.1242/jeb.151.1.57>.
- Ballintijn, C., 1972. Efficiency, mechanics and motor control of fish respiration. Respir. Physiol. 14, 125–141. URL: [https://doi.org/10.1016/0034-5687\(72\)90023-0](https://doi.org/10.1016/0034-5687(72)90023-0).
- Barreto, R.E., Volpato, G.L., 2011. Ventilation rates indicate stress-coping styles in nile tilapia. J. Biosci. 36, 851–855. URL: <https://doi.org/10.1007/s12038-011-9111-4>.
- Bowden, A., Adams, M., Andrewartha, S., Elliott, N., Frappell, P., Clark, T., 2022. Amoebic gill disease increases energy requirements and decreases hypoxia tolerance in Atlantic salmon (*Salmo Salar*) smolts. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 265, 111128. URL: <https://doi.org/10.1016/j.cbpa.2021.111128>.
- Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A., 2020. Albumentations: fast and flexible image augmentations. Information 11, 125. URL: <https://doi.org/10.3390/info11020125>.
- Cisar, P., Bekkozhayeva, D., Movchan, O., Saberoon, M., Schraml, R., 2021. Computer vision based individual fish identification using skin dot pattern. Sci. Rep. 11, 16904. URL: <https://doi.org/10.1038/s41598-021-96476-4>.

- Crouse, D.F., 2016. On implementing 2d rectangular assignment algorithms. *IEEE Trans. Aerosp. Electron. Syst.* 52, 1679–1696. URL: <https://doi.org/10.1109/TAES.2016.140952>.
- Davis, L.E., Schreck, C.B., 1997. The energetic response to handling stress in juvenile coho salmon. *Trans. Am. Fish. Soc.* 126, 248–258. URL: [10.1577/1548-8659\(1997126-0248:TERTHS>2.3.CO;2](https://doi.org/10.1577/1548-8659(1997126-0248:TERTHS>2.3.CO;2).
- Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 248–255. URL: <https://doi.org/10.1109/CVPR.2009.5206848>.
- Ellis, T., James, J.D., Stewart, C., Scott, A.P., 2004. A non-invasive stress assay based upon measurement of free cortisol released into the water by rainbow trout. *J. Fish Biol.* 65, 1233–1252. URL: <https://doi.org/10.1111/j.0022-1112.2004.00499.x>.
- Føre, M., Alver, M.O., Alfredsen, J.A., Senneset, G., Espmark, Åsa, Terjesen, B.F., 2018. Modelling how the physical scale of experimental tanks affects salmon growth performance. *Aquaculture* 495, 731–737. URL: <https://doi.org/10.1016/j.aqua.2018.06.057>.
- Fry, J.P., Mailloux, N.A., Love, D.C., Milli, M.C., Cao, L., 2018. Feed conversion efficiency in aquaculture: do we measure it correctly? *Environ. Res. Lett.* 13, 024017. URL: <https://doi.org/10.1088/1748-9326/aaa273>.
- Gómez-Vargas, N., Alonso-Fernández, A., Blanquero, R., Antelo, L.T., 2023. Re-identification of fish individuals of undulate skate via deep learning within a few-shot context. *Eco. Inform.* 75, 102036. URL: <https://doi.org/10.1016/j.ecoinf.2023.102036>.
- Hartley, R.I., Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*, Second ed. Cambridge University Press.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 770–778. URL: <https://doi.org/10.1109/CVPR.2016.90>.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 2980–2988. URL: <https://doi.org/10.1109/ICCV.2017.322>.
- Israeli, D., Kimmel, E., 1996. Monitoring the behavior of hypoxia-stressed carassius auratus using computer vision. *Aquac. Eng.* 15, 423–440. URL: [https://doi.org/10.1016/S0144-8609\(96\)01009-6](https://doi.org/10.1016/S0144-8609(96)01009-6).
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82, 35–45. URL: <https://doi.org/10.1115/1.3662552>.
- Kämmer, N., Reimann, T., Ovcharova, V., Braunbeck, T., 2023. A novel automated method for the simultaneous detection of breathing frequency and amplitude in zebrafish (*danio rerio*) embryos and larvae. *Aquat. Toxicol.* 258, 106493. URL: <https://doi.org/10.1016/j.aquatox.2023.106493>.
- Kinkead, R., Perry, S.F., 1991. The effects of catecholamines on ventilation in rainbow trout during hypoxia or hypercapnia. *Respir. Physiol.* 84, 77–92. URL: [https://doi.org/10.1016/0034-5687\(91\)90020-J](https://doi.org/10.1016/0034-5687(91)90020-J).
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, 25. Curran Associates, Inc., pp. 1097–1105. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf)
- Li, M., Huang, Y., Gong, S.P., 2019. Assessing the risk of potentially hazardous asteroids through mean motion resonances analyses. *Astrophys. Space Sci.* 364, 78. URL: <https://doi.org/10.1007/s10509-019-3557-5>.
- Liu, Z., Li, X., Fan, L., Lu, H., Liu, L., Liu, Y., 2014. Measuring feeding activity of fish in ras using computer vision. *Aquac. Eng.* 60, 20–27. URL: <https://doi.org/10.1016/j.aquaeng.2014.03.005>.
- Lourakis, M.I., 2005. A brief description of the levenberg-marquardt algorithm implemented by levmar. *Foundat. Res. Technol.* 4, 1–6. URL: [https://www.researchgate.net/publication/239328019\\_A\\_Brief\\_Description\\_of\\_the\\_Levenberg-Marquardt\\_Algorithm\\_Implemented\\_by\\_levmar](https://www.researchgate.net/publication/239328019_A_Brief_Description_of_the_Levenberg-Marquardt_Algorithm_Implemented_by_levmar).
- Magid, A.M.A., Babiker, M.M., 1975. Oxygen consumption and respiratory behaviour of three nile fishes. *Hydrobiologia* 46, 359–367. URL: <https://doi.org/10.1007/BF00028279>.
- Martins, C.I.M., Galhardo, L., Noble, C., Damsgård, B., Spedicato, M.T., Zupa, W., Beauchaud, M., Kulczykowska, E., Massabuau, J.C., Carter, T., Planellas, S.R., Kristiansen, T., 2012. Behavioural indicators of welfare in farmed fish. *Fish Physiol. Biochem.* 38, 17–41. URL: <https://doi.org/10.1009/s10695-011-9518-8>.
- Mathisen, B.M., Bach, K., Meidell, E., Måløy, H., Sjøblom, E.S., 2020. Fishnet: A unified embedding for salmon recognition. In: 24th European Conference on Artificial Intelligence, pp. 3001–3008. URL: <https://doi.org/10.3233/FAIA200475>.
- Mauricio, J., Domingues, I., Bernardino, J., 2023. Comparing vision transformers and convolutional neural networks for image classification: a literature review. *Appl. Sci.* 13 <https://doi.org/10.3390/app13095521>. URL: <https://doi.org/10.3390/app13095521>.
- McCormick, S., Shrimpton, J., Carey, J., O'Dea, M., Sloan, K., Moriyama, S., Björnsson, B., 1998. Repeated acute stress reduces growth rate of Atlantic salmon parr and alters plasma levels of growth hormone, insulin-like growth factor i and cortisol. *Aquaculture* 168, 221–235. URL: [https://doi.org/10.1016/S0044-8486\(98\)00351-2](https://doi.org/10.1016/S0044-8486(98)00351-2).
- Nilsson, J., Gismervik, K., Nielsen, K.V., Iversen, M.H., Noble, C., Kolarevic, J., Frotjold, H., Nilsen, K., Wilkinson, E., Klakegg, B., et al., 2022. Lakselv-standardisert operasjonell velferdsovervåking for laks i matfiskanlegg. Technical Report, Institute of Marine Research. URL: <https://hdl.handle.net/11250/2997021>.
- Noble, C., Gismervik, K., Iversen, M., Kolarevic, J., Nilsson, J., Stien, L., Turnbull, J., 2018. Welfare Indicators for Farmed Atlantic Salmon: Tools for Assessing Fish Welfare. Technical Report, Nofima. URL: <https://hdl.handle.net/11250/2575780>.
- Nygård, T.A., Jahren, J.H., Schellewald, C., Stahl, A., 2022. Motion trajectory estimation of salmon using stereo vision. *IFAC-PapersOnLine* 55, 363–368.
- Oldham, T., Oppedal, F., Dempster, T., 2018. Cage size affects dissolved oxygen distribution in salmon aquaculture. *Aquaculture environment. Interactions* 149–156. URL: <https://doi.org/10.3354/aei00263>.
- Olsen, R., Sundell, K., Hansen, T., Hemre, G.I., Myklebust, R., Mayhew, T., Ringoe, E., 2002. Acute stress alters the intestinal lining of Atlantic salmon, *salmo Salar* L.: An electron microscopical study. *Fish Physiol. Biochem.* 26, 211–221. URL: <https://doi.org/10.1023/A:1026217719534>.
- Øvrebo, T.K., Balseiro, P., Imsland, A.K.D., Stefansson, S.O., Tvetenås, R., Sveier, H., Handeland, S., 2022. Investigation of growth performance of post-smolt Atlantic salmon (*salmo Salar* L.) in semi closed containment system: a big-scale benchmark study. *Aquac. Res.* 53, 4178–4189. URL: <https://doi.org/10.1111/are.15919>.
- Pandurangan, M., Moorthy, H., Sambandam, R., Jeyaraman, V., Irisappan, G., Kothandam, R., 2014. Effects of stress hormone cortisol on the mRNA expression of myogenin, myoD, myf5, pax3 and pax7. *Cytotechnology* 66, 839–844. URL: <https://doi.org/10.1007/s10616-013-9635-6>.
- Parra, L., Lloret, G., Lloret, J., Rodilla, M., 2018. Physical sensors for precision aquaculture: a review. *IEEE Sensors J.* 18, 3915–3923. URL: <https://doi.org/10.1109/JSEN.2018.2817158>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. Pytorch: An imperative style, high performance deep learning library. In: Advances in Neural Information Processing Systems, 32. Curran Associates, Inc., pp. 8024–8035. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2019/file/bdbca288fee7f92f2bfaf97012727740-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfaf97012727740-Paper.pdf)
- Petersen, J., Bracke, M., Midtlyng, P.J., Folkedal, O., Stien, L., Steffenak, H., Kristiansen, T., 2013. Salmon welfare index model 2.0: An extended model for overall welfare assessment of caged Atlantic salmon, based on a review of selected welfare indicators and intended for fish health professionals. *Rev. Aquac.* 6, 162–179. URL: <https://doi.org/10.1111/raq.12039>.
- Pinkiewicz, T., Purser, G., Williams, R., 2011. A computer vision system to analyse the swimming behaviour of farmed fish in commercial aquaculture facilities: a case study using cage-held Atlantic salmon. *Aquac. Eng.* 45, 20–27. URL: <https://doi.org/10.1016/j.aquaeng.2011.05.002>.
- PME, 2021. Minidot Logger User's Manual. URL: <https://www.pme.com/products/minidot>.
- Raguram, R., Chum, O., Pollefeys, M., Matas, J., Frahm, J.M., 2013. Usac: a universal framework for random sample consensus. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 2022–2038. URL: <https://doi.org/10.1109/TPAMI.2012.257>.
- Randall, D.J., Taylor, E., 1991. Evidence of a role for catecholamines in the control of breathing in fish. *Rev. Fish Biol. Fish.* 1, 139–157. URL: <https://doi.org/10.1007/BF00157582>.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788. URL: <https://doi.org/10.1109/CVPR.2016.91>.
- Remen, M., 2012. The Oxygen Requirement of Atlantic Salmon (*Salmo salar* L.) in the on-Growing Phase in Sea Cages. Phd thesis. University of, Bergen. URL: <https://hdl.handle.net/1956/9535>.
- Schellewald, C., Saad, A., Stahl, A., 2024. Mouth opening frequency of salmon from underwater video exploiting computer vision. In: IFAC-PapersOnLine. Accepted.
- Solstrom, D., Oldham, T., Solstrom, F., Klebert, P., Stien, L.H., Vågseth, T., Oppedal, F., 2018. Dissolved oxygen variability in a commercial sea-cage exposes farmed Atlantic salmon to growth limiting conditions. *Aquaculture* 486, 122–129. URL: <https://doi.org/10.1016/j.aqua.2017.12.008>.
- Sopinka, N.M., Donaldson, M.R., O'Connor, C.M., Suski, C.D., Cooke, S.J., 2016. 11 - stress indicators in fish. In: *Biology of Stress in Fish*, 35. Academic Press, pp. 405–462 of fish physiology. URL: <https://doi.org/10.1016/B978-0-12-802728-8.00011-4>.
- Stien, L.H., Nilsson, J., Bui, S., Fosseidengen, J.E., Kristiansen, T.S., Øverli, Ø., Folkedal, O., 2017. Consistent melanophore spot patterns allow long-term individual recognition of Atlantic salmon *salmo Salar*. *J. Fish Biol.* 91, 1699–1712. URL: <https://doi.org/10.1111/jfb.13491>.
- Tan, M., Le, Q., 2019. EfficientNet: rethinking model scaling for convolutional neural networks. In: Proceedings of the 36th International Conference on Machine Learning. PMLR, pp. 6105–6114. URL: <https://proceedings.mlr.press/v97/tan19a.html>.
- Taylor, E., 2011. Control of respiration – Generation of the respiratory rhythm in fish. In: *Encyclopedia of Fish Physiology*. Academic Press, pp. 854–864. URL: <https://doi.org/10.1016/B978-0-12-374553-8.00250-1>.
- Torrisen, O., Olsen, R., Toresen, R., Hemre, G.I., Tacon, A., Asche, F., Hardy, R., Lall, S., 2011. Atlantic salmon (*salmo Salar*): the “super-chicken” of the sea? *Rev. Fish. Sci.* 19, 257–278. URL: <https://doi.org/10.1080/10641262.2011.597890>.
- Wade, N.M., Clark, T.D., Maynard, B.T., Atherton, S., Wilkinson, R.J., Smullen, R.P., Taylor, R.S., 2019. Effects of an unprecedented summer heatwave on the growth performance, flesh colour and plasma biochemistry of marine cage-farmed Atlantic salmon (*salmo Salar*). *J. Therm. Biol.* 80, 64–74. URL: <https://doi.org/10.1016/j.jthermobio.2018.12.021>.
- Wei, Y., Jiao, Y., An, D., Li, D., Li, W., Wei, Q., 2019. Review of dissolved oxygen detection technology: from laboratory analysis to online intelligent detection. *Sensors* 19, 3995. URL: <https://doi.org/10.3390/s19183995>.

- Wendelaar, B.S., 1997. The stress response in fish. *Physiol. Rev.* 77, 591–625. URL: <https://doi.org/10.1152/physrev.1997.77.3.591>.
- Zheng, H., Liu, R., Zhang, R., Hu, Y., 2014. A method for real-time measurement of respiratory rhythms in medaka (*Oryzias latipes*) using computer vision for water quality monitoring. *Ecotoxicol. Environ. Saf.* 100, 76–86. URL: <https://doi.org/10.1016/j.ecoenv.2013.11.016>.
- Zhou, Z., Hitt, N.P., Letcher, B.H., Shi, W., Li, S., 2022. Pigmentation-based visual learning for *Salvelinus fontinalis* individual re-identification. In: 2022 IEEE International Conference on Big Data (Big Data), pp. 6850–6852. URL: <https://doi.org/10.1109/BigData55660.2022.10020966>.