IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 May 2024, accepted 26 May 2024, date of publication 4 June 2024, date of current version 13 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3409360

RESEARCH ARTICLE

Deep Learning for Automated Egg Maturation Prediction of Atlantic Salmon Using Ultrasound Imaging

YASIN YARI[1], (Graduate Student Member, IEEE), INGUN NÆVE[2], PER HELGE BERGTUN[3], ASLE HAMMERDAL[4], SVEIN-ERIK MÅSØY[1], (Member, IEEE), MARCO MARIEN VOORMOLEN[5], AND LASSE LØVSTAKKEN[1], (Member, IEEE)

[1]Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway
[2]Lerøy Seafood Group, 5006 Bergen, Norway
[3]MOWI AS, 7044 Trondheim, Norway
[4]AquaGen AS, 7010 Trondheim, Norway
[5]InPhase Solutions AS, 7038 Trondheim, Norway

Corresponding author: Yasin Yari (yasin.yari@ntnu.no)

This work was supported by the Research Council of Norway facilitated by the Center for Innovative Ultrasound Solutions (CIUS) hosted by the Department of Circulation and Medical Imaging (ISB), Norwegian University for Science and Technology (NTNU).

ABSTRACT The Atlantic salmon maturation process has been studied for decades to increase the quantity and quality of the production in farming facilities. An important topic in this context is the salmon egg maturation process. Ultrasound imaging is considered an effective tool for monitoring the egg development stage of salmon, but manual inspection is time-consuming and dependent on operator experience. We propose a method for automated monitoring of the egg maturation stage in salmon using deep learning, providing complimentary decisions on egg morphology. A segmentation network was developed to solve the challenge of separating and measuring individual eggs in the ovary. The segmentation part was combined with a classification network to determine the maturation stage of the eggs. Our model was able to segment eggs and classify their development stage with over 88% accuracy, outperforming established methods designed for similar tasks. A real-time application was developed which provided an estimation of size and maturity stage while scanning. The egg state estimation showed potential for replacing manual evaluations and can enable fully automatic evaluation of maturation in Atlantic salmon.

INDEX TERMS Deep learning, ultrasound, maturation monitoring, salmon maturation, egg maturation prediction.

I. INTRODUCTION

Atlantic salmon farming stands out as one of the most significant sectors within the field of aquaculture. Breeding companies focus on supplying high-quality salmon eggs by meticulously replicating the natural aquatic environment of salmon in both sea and land-based facilities. The salmon which are utilized for egg production have undergone genetic selection over several decades to enhance their growth and overall productivity, making them highly valuable assets [1]. To ensure a year-round supply of eggs for consistent production and optimal facility utilization, it is important to have matured salmon available at all times. Since salmon eggs cannot be preserved for an extended period, controlling the maturation process becomes imperative. Shortening or lengthening the maturation period can be used to harvest fresh eggs from salmon at the time requested by the customer. This brings out the need to continuously monitor the maturation stage of salmon and has led to the development of various methods for maturation monitoring.

The associate editor coordinating the review of this manuscript and approving it for publication was Binit Lukose.

Among several invasive methods, ultrasound imaging is an approved and non-invasive method for monitoring salmon maturation, and for identifying egg development stages [2], [3]. The ultrasound assessment of salmon for maturation monitoring has two primary stages; evaluation of ovary size development from early maturation to the stripping time, and evaluation of the egg morphology from a point where eggs are sufficiently visible in ultrasound images (approximately 6-8 weeks after transferring to fresh water) up until the stripping time. Currently, the evaluation of egg morphology is a manual process, where experts examine fish eggs within the ovary and classify them into four main development classes, as illustrated in Figure 1.

Due to the labor-intensive and subjective nature of the manual ultrasound evaluation process, it is common practice to assess only a limited population of fish within a tank and extrapolate the results to the entire brood stock. Additionally, this evaluation can be subject to variability, particularly during the early and final phases of maturation. By automating the ultrasound maturation grading process, we can eliminate operator dependency and reduce stress on the fish. Additionally, the automation facilitates the logging of more comprehensive data about salmon eggs, including details such as their size, biological attributes, egg quality (which can influence the survival, performance, health, and welfare of the offspring), and overall condition within each individual salmon. This information is valuable for the fish farming facility and provides important data for future research.

To automate this procedure, salmon eggs should be identified and separated and their stage of maturation should be estimated, which can be based on image segmentation and classification respectively. We approach this using deep learning-based image analysis, which has proven to be effective and accurate for a number of similar tasks. A particular challenge in this context is the compact arrangement of eggs within the ovary, complicating the segmentation of eggs as distinct entities.

To investigate the literature for potential solutions for the application-specific challenges, we conducted a literature survey with an emphasis on the segmentation of closely situated or small objects. We identified potential solutions, including the utilization of Watershed, the creation of Horizontal/Vertical maskings, diverse feature extractions, multi-thresholding, and the incorporation of Attention mechanisms, tailored to address the unique aspects of our problem.

The watershed method is proven useful in avoiding over-segmentation situations where the objects are very close to each other [4]. It has been used for different segmentation applications such as partitioning 3D surface meshes [5], flaw detection in radiographic weldment images [5], randomly textured color image segmentation [6], and in ultrasound technology for breast tumor segmentation [7]. Zhang et. al. [8] introduced a Watershed image segmentation algorithm for bubble segmentation. Study [9] employs a multi-thresholding strategy to detect compact objects in images, integrating geometric characteristics into the decision-making process. Study [10] tackles the segmentation of closely positioned objects through a two-stage instance segmentation method. This approach creates a center-aware feature representation for predicting masks. In the first stage, objects are segmented against the background, and their centers are determined. Then, a position-based encoding is utilized to combine both geometric and semantic features for further refinement.

Notably, Attention-based approaches for small object segmentation have been investigated in the studies conducted by Fei et al. [11], Sang et al. [12], and Zhang et al. [13]. This methodology holds the potential to enhance the accuracy of segmenting a greater number of eggs within the ovary.

Graham et al. [14] introduced the Hover-Net method mainly for segmenting nuclei in histology images. Hover-Net has been proven accurate while avoiding intersections in overlapping regions of clustered nuclei by creating horizontal and vertical feature maps. The network detects the objects(nucleus) type in the up-sampling classification which is made alongside the segmentation (H/V maps) using the shared encoder. These maps were then used to separate closely situated objects in a postprocessing step. This proposed method is able to segment and classify compact objects.

These various segmentation approaches from the literature provide valuable insights and techniques that could be adapted and explored for the specific task of automating the segmentation of salmon eggs. This work's primary contribution is the use of deep learning and ultrasound technology to autonomously predict the maturation state of Atlantic salmon eggs. The approach involves data sampling, model development, model evaluation, and an initial real-time application test in a live-field setting. The developed application demonstrates efficient segmentation of compact eggs within the ovary and provides accurate assessments of their developmental stage.

II. DATA COLLECTION AND ANNOTATION

During this study, a dataset of Atlantic salmon ovaries was recorded in two time frames of May to September and June to December. Each survey consisted of 100 PIT-tagged salmon in brood-stock facilities located in Kyrksæterøra (AquaGen) and Lysøysundet (MOWI) in Norway. While the observation period encompasses the entire duration from the transfer of fish to the land facility up to the stripping time, the primary dataset for this study predominantly originates from the period from eight weeks after transfer to freshwater when most of the eggs become visible in ultrasound images. For additional details regarding the fish involved in this study, sampling protocol, and rearing conditions, please refer to [2] and [15].

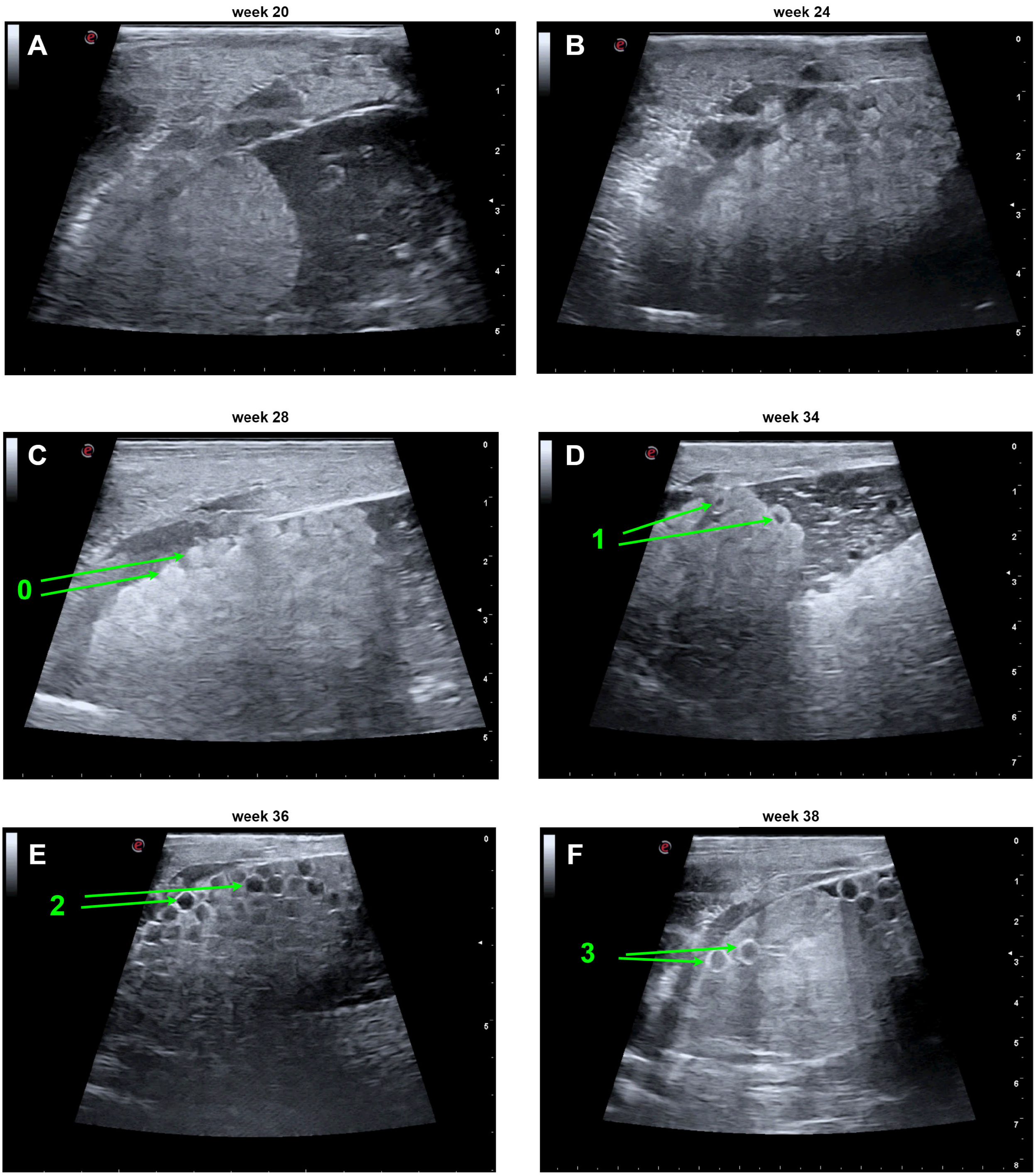

FIGURE 1. **Various stages of the maturation of salmon eggs (four main egg morphology stages is shown in green); Two weeks after transferring fish to freshwater where eggs are mostly not visible (A), Six weeks after the transfer to freshwater where very tiny eggs are visible in some salmon (B), 10 weeks after transfer to freshwater where eggs are visible on most fish and considered as class=0 of egg maturation (C), 15 weeks after landing fish to freshwater where a small dark spot is visible on some eggs representing class=1 of the egg development (D), 17 weeks after transfer to freshwater the harvesting period and the dark spot has covered almost the entire eggs which represent egg maturation class=2 where the dark spot has taken more than 50% of the egg area, 19 weeks into the maturation and eggs are released and floating within the abdominal cavity and the liver has become visible which is class=3 stating the eggs are ready for spawning (F). In this work, for clarity and simplicity, we use the term 'egg' to refer to both the oocyte (before ovulation) and the egg (after ovulation). While the terms 'oocyte' and 'egg' are commonly used to distinguish pre- and post-ovulation stages, we adopt a unified terminology and consistently refer to them as 'egg' throughout this article.**

The data has been annotated via the open-source tool AnnotationWeb [16], using a spline segmentation tool. Additionally, we have added functionality to the software to extract COCO data formats [17] for instance-wise evaluation

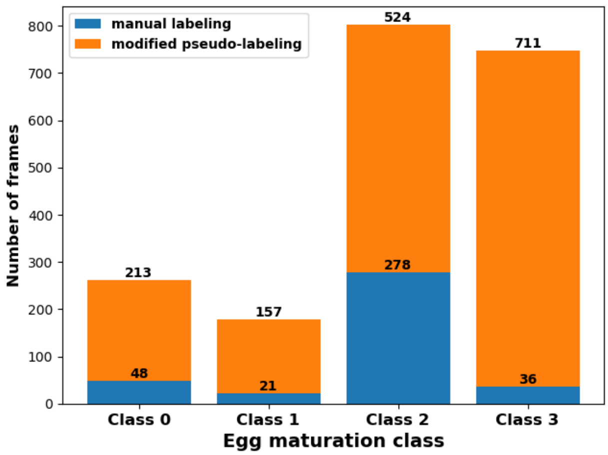

FIGURE 2. Overview of egg morphology dataset.

of individual eggs and avoiding the intersection of the segmented areas. The data has initially been annotated manually and further by manual supervising and modification of the Pseudo-labeling technique in [18]. Figure 2 provides an overview of the dataset. The dataset exhibits a class imbalance, primarily attributed to its structure. The imbalance arises due to the brief duration of class 1 egg development, while the segmentation of eggs in class 0 is challenging, given their early and hard-to-segment stages.

III. METHOD

A. NETWORK

Figure 3 provides a visual representation of the architecture employed in our application. Our primary objective is to segment compact eggs within the ovary as individual entities. The automation of this application can be achieved through a two-step process involving classification to determine the egg state and segmentation to measure egg size. Alternatively, a more comprehensive approach involves multi-class segmentation to simultaneously identify both egg state and size. However, segmenting closely packed eggs that appear to overlap presents a significant challenge. When using semantic segmentation, egg masks may intersect, leading to potential overestimation of egg sizes. While the instance segmentation method offers a solution to this issue, it's worth noting that it may struggle to detect the boundaries of closely located objects.

Using a set of convolution blocks via a targeted transfer learning for down-sampling is proven to be effective and time efficient when the architecture is not very deep [19]. Our approach uses a shared feature extraction via a set of convolutions blocks obtained from a U-Net used in Leclerc, and Smistad et al. [19]. It outputs a classification head using one flattening and one dense layer, with a linear activation function at the lowest down-sampling level. The full segmentation model is depicted in Figure 3. The utilization of the linear activation function enables the adjustment of prediction thresholds, thereby reducing the occurrence of false positives. Additionally, it enhances the interpretability of the model, as the weights in a linear layer directly represent the significance or contribution of individual input features to the output. The model also up-samples the shared latent space into a binary segmentation (all egg types vs background), together with horizontal and vertical maps of the objects in a 4-channel output. The horizontal and vertical maps are built to produce high values when intersecting together, making them detectable. Thus, we replaced the multi-class segmentation approach for four types of eggs with binary segmentation, allowing the classification head to determine the stage of the eggs (4-class) based on full image information and not individual egg annotations. Moreover, the classification head includes an additional output specifically designed to detect the presence of eggs, resulting in a classification system with 4+1 classes. This extension aims to reduce the occurrence of false positive predictions. When the frame does not contain any eggs, a designated background class is utilized to adjust the segmentation head, thereby mitigating the risk of erroneous positive predictions

In the case of having two classes at the same frame, we consider the highest class as the egg maturation level as it has been the common maturation grading procedure. Therefore, there is no need for patch-wise classification or up-sampling for the classification head in our application. As for the segmentation head, the up-sampling architecture uses each level of the encoder via concatenation to provide the output.

The outputs of each head of the network are then post-processed to create the final segmentation mask.

One of the main benefits of the presented architecture is the advantage of multi-task learning via shared parameters. It can help the model to focus on the general features that are common for each task rather than task-specific features [20]. To consider the Attention effect, mainly for segmenting more eggs by focusing on deeper regions of the ovary in the image and to generate better descriptions for those deeper ovary regions, an Attention gate has been added to the up-sampling levels same as the architecture given in [21]. We explored various segmentation techniques to enhance the detection of eggs within images. Firstly, we employed the Attention mechanism within our architecture. This mechanism dynamically prioritizes features extracted by the encoder, assigning importance to them based on their relevance to segmentation. By computing Attention scores between all pairs of positions in feature maps, we can selectively focus on significant spatial locations, improving the accuracy of segmentation.

Additionally, we experimented with the watershed algorithm. This approach treats pixel intensities as a topographic surface and mimics a flooding process from specified markers to divide the image into distinct regions [22]. As flooding progresses, valleys fill up to form basins, and boundaries between basins, known as watershed lines, delineate object boundaries [23]. This technique offers a unique perspective

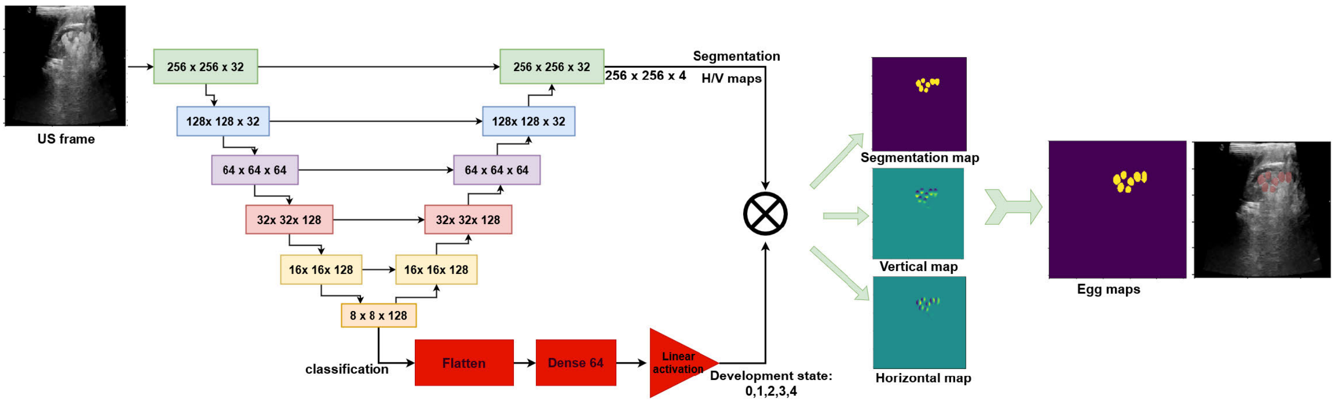

FIGURE 3. Overview of egg segmentation model. A modified U-Net is employed to deliver both classification and segmentation outputs. The segmentation output includes horizontal and vertical map data, as well as a segmentation mask essential for egg boundary identification in a postprocessing step.

on segmentation by leveraging the spatial relationships within the image.

Finally, we investigated instance segmentation using YOLACT. Unlike traditional segmentation methods, instance segmentation architectures like YOLACT, based on concepts such as Mask R-CNN, employ sophisticated features such as Feature Pyramid Network (FPN) and Region Proposal Network (RPN). These components enable the analysis of objects at different scales and suggest regions of interest. The mask prediction head generates segmentation masks for individual objects, refining detections through techniques like Non-Maximum Suppression for contextual understanding. By considering multi-scale features and proposing regions, YOLACT may segment objects with high precision and speed [24].

B. PRE-PROCESSING- CREATING H/V FEATURE MAPS

To generate the horizontal and vertical maps, the distance of each channel of input image pixels to the center of the object (in this case, eggs) is calculated. This process results in maps that show how far each pixel is from the object center along both horizontal and vertical axes. The pixel values associated with the object are scaled within the range of $[-1, 1]$, based on the maximum absolute value observed in each map. This scaling ensures that the center point of each egg corresponds to a value of zero.

C. POST-PROCESSING- GENERATING FINAL EGG MASKS

In order to create the final output egg masks from the Horizontal (H) and Vertical (V) maps in conjunction with the segmentation head, a series of steps are performed. Figure 4 provides a flow chart for egg mask creation. This procedure involves determining the magnitude of eggs within a frame by analyzing the H/V maps. This magnitude is derived by taking the square root of the sum of the squared values obtained from the H/V maps. Furthermore, we employ Sobel filters to detect both the horizontal and vertical edges of the eggs. Subsequently, we calculate the maximum value (element-wise) between the horizontal and vertical edges, effectively merging responses from both directions and yielding a single-edge response value for each pixel. The resulting tensor is then multiplied by the segmentation head tensor, to fuse prediction information with a mask. This step helps mitigate potential mismatches in predictions, particularly in the detection of eggs in the lower parts of the ovary where the prediction confidence within each egg pixel may vary.

The generated egg mask is combined with classification prediction to detect the stage of the eggs. The classification head generates a 5-class output, where each class represents a specific stage of egg development. These classes are aligned with the channel axis: channel 0 indicates the absence of eggs, while channels 1 through 4 signify different egg developmental stages. Initially, we determine the channel number associated with the maximum value across the tensor's axes. Subsequently, this channel number, representing the most probable egg stage, is employed as a scalar value to scale each element of the segmentation output via straightforward element-wise multiplication

The whole post-processing steps were combined with the network as a separate custom layer to enable a direct end-to-end prediction.

IV. IMPLEMENTATION

We implemented our deep learning model using TensorFlow 2.6.0 on a laptop equipped with an Intel(R) Core(TM) i7-10750H CPU operating at 2.60GHz, 32GB of RAM, and an NVIDIA RTX 2080 GPU with 8GB of dedicated memory.

A. NETWORK TRAINING SETUP

The network is composed of a segmentation head with four channels, encompassing segmentation and H/V maps, along with an extra classification head. In the computation of the overall loss function, a weighted cross-entropy loss is implemented for the classification head. Meanwhile, a blend of Dice and mean square error loss is utilized for the segmentation and H/V maps to account for both discrete and continuous variables in the output head. The total loss is balanced with a 10:1 ratio, favoring the segmentation and H/V

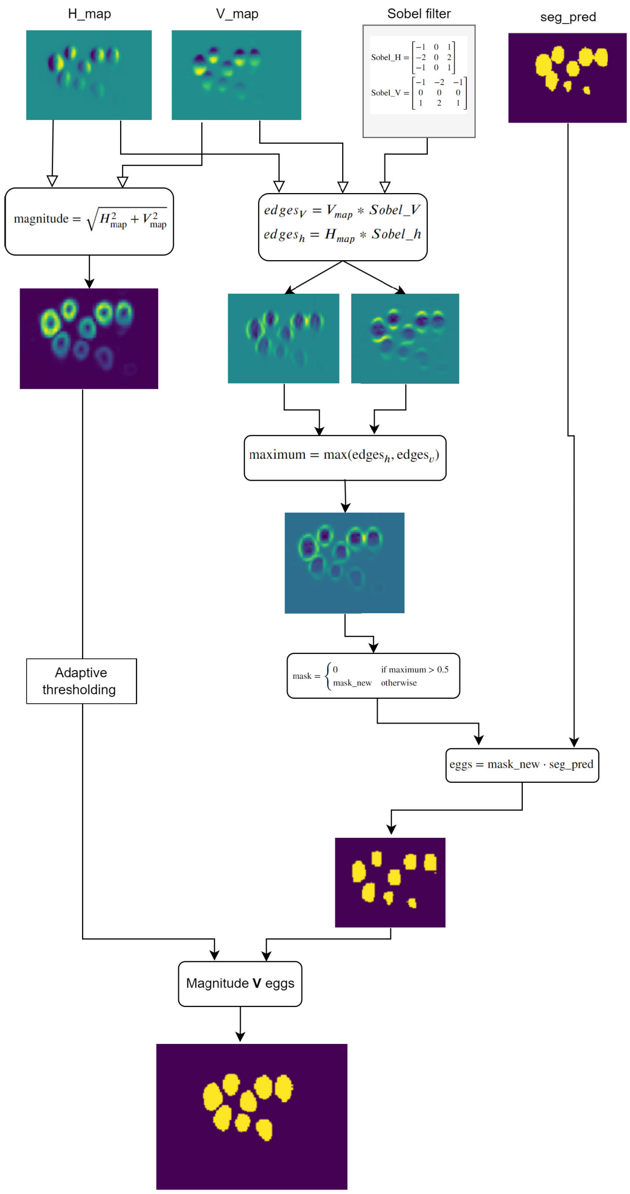

FIGURE 4. The post-processing steps to create egg masks.

maps. This weighting emphasizes the importance of accurate segmentation and mapping in the overall training process. Our data augmentations include random rotations within a range of 30 degrees, the introduction of random noise, gamma transformations, elastic deformations, Gaussian shadowing, and image blurring. The model was trained using Adam optimizer for 50 epochs using mini-batch gradient descent where 20% of the training set is used for the validation. The dataset was split into 85% for training and 15% for testing.

B. EGG SIZE AND STAGE ESTIMATION

The final segmentation model was used to predict the egg size on the test dataset. The top 20% largest segmented eggs are obtained from each fish. This is mainly to cover the possibility that the 2D ultrasound cross-section may sometimes cross a secant segment of an egg rather than its diameter. To predict the egg stage, we used a high threshold (>6) for our classification head. This threshold was primarily set to ensure more robust predictions. If the prediction score falls below this threshold, we only extract the egg size, while the stage is determined based on other frames in the head-to-tail scan.

TABLE 1. Binary segmentation results.

Methods	Dice Score	Hausdorff
U-Net	0.878 $\pm$0.05	4.54 $\pm$1.15
U-Net with Attention	0.902 $\pm$0.09	3.89 $\pm$1.5
U-Net with Watershed	0.792 $\pm$0.12	8.74 $\pm$4.5
Instance segmentation (YOLACT)	0.865 $\pm$0.09	7.9 $\pm$2.5
Proposed model	0.891 $\pm$0.09	4.02 $\pm$1.6

C. REAL-TIME APPLICATION

A real-time application was implemented using the FAST framework [25], [26], which includes GPU processing and visualization, and several deep learning inference engines. Data was streamed in real-time from the ultrasound machine based on frame grabbing. Further real-time processing included pre-processing (cropping, resizing, and normalization), rendering, and post-processing the egg data to obtain egg volume. The application with the proposed network ran at 249 frames per second on GPU and 16 frames per second on CPU, enabling real-time feedback during imaging.

V. RESULTS

The evaluation result of the segmentation models is provided for both binary and multi-class segmentation in Table 1 and 2. We first performed a binary segmentation task where we segmented eggs of any type and stage to provide an initial segmentation performance benchmark of the models for this specific task. As shown in Table 1, the binary segmentation, segments egg vs background where the models are baseline U-Net [19], baseline U-Net with Attention [21], baseline U-Net with watershed [27], instance segmentation using YOLACT [24], and our proposed method.

Then a multi-class segmentation performance comparison was carried out in Table 2 to assign a label (maturation stage) to the segmented eggs. Furthermore, the classification results of the egg development stage prediction are provided. The classification performance was assessed by comparing the segmentation prediction classes to the reference. In the case of our proposed model, we evaluated the performance of the classification output.

The model selection was made to compare a robust baseline with additional functionalities such as the proposed changes mentioned in the method section, Attention, and Watershed. The YOLACT segmentation network [24] was included to compare the performance of a real-time instance segmentation model as well. The outcomes of training the models until convergence and subsequent testing them show that our proposed modifications effectively maintain Dice and Hausdorff metrics when transitioning from binary to

TABLE 2. The result of the multi-class segmentation and classification of the eggs.

Methods	Segmentation		Classification		
	Dice Score	Hausdorff	Precision	Recall	F1 score
U-Net	0.853 ±0.16	4.51 ±1.6	0.757 ±0.42	0.679 ±0.42	0.704 ±0.41
U-Net with Attention	0.821 ±0.16	4.24 ±1.5	**0.948 ±0.21**	0.795 ±0.29	0.844 ±0.25
Instance segmentation (YOLACT)	0.710 ±0.05	5.5 ±1.8	0.740 ±0.25	0.815 ±0.24	0.775 ±0.38
Proposed model (with Attention)	0.851 ±0.17	4.48±1.6	0.800 ±0.39	0.802 ±0.39	0.801 ±0.39
Proposed model	**0.894 ±0.14**	**4.11 ±1.6**	0.885 ±0.31	**0.886 ±0.31**	**0.885 ±0.31**

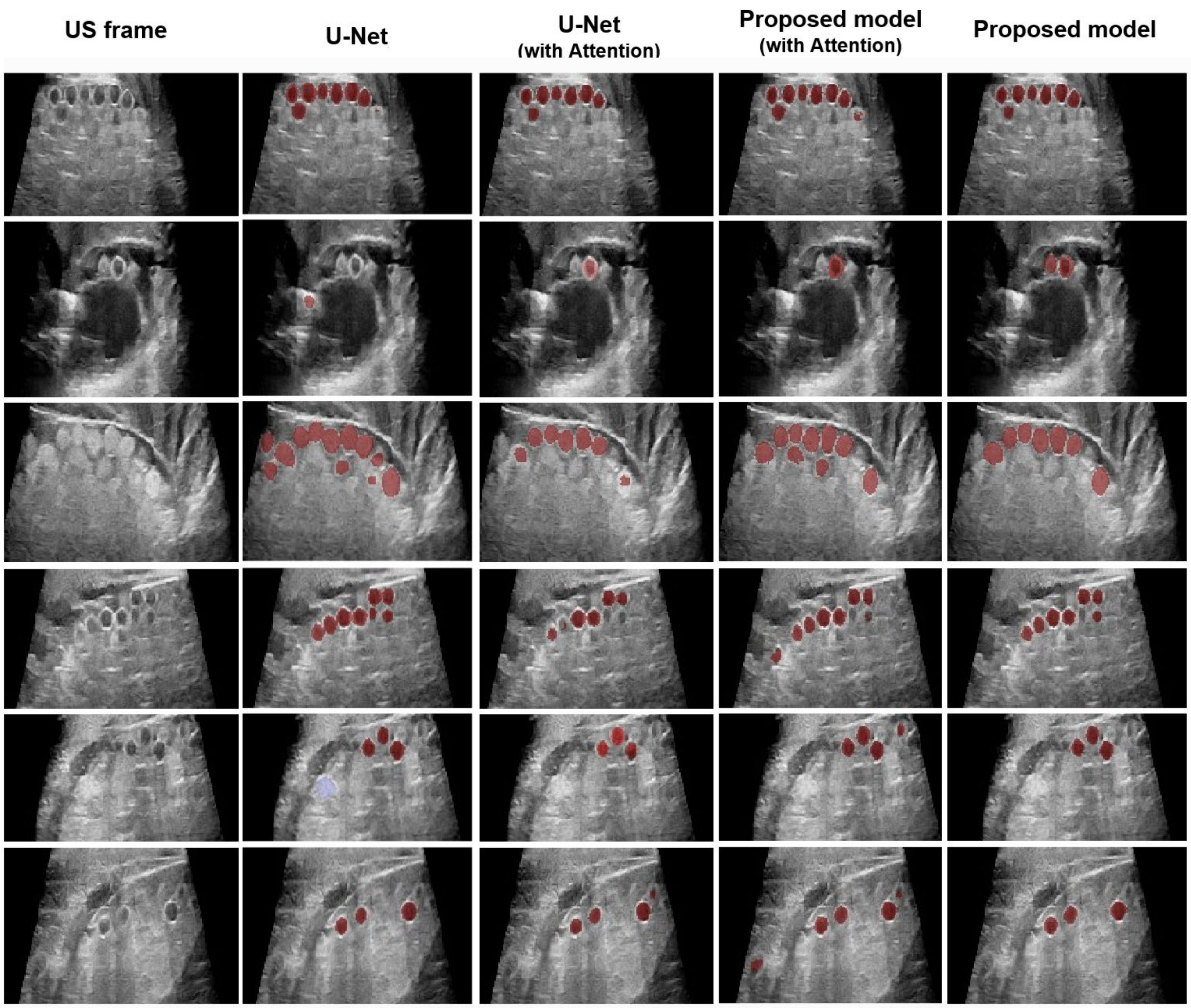

FIGURE 5. Segmentation quality in different models.

multi-class segmentation. The instance segmentation model specifically loses precision when it is used for multi-class egg segmentation. The main cause of this performance drop seems to be poor egg stage classification rather than segmentation score when comparing segmentation and classification performance. The same trend is happening for the baseline U-Net as well which is expected as the multi-class segmentation task is more complex. The proposed model was created with and without Attention. In the proposed model with Attention, the multi-class performance did not improve the results in our work. The classification performance drop is also considerable when using Attention.

The example results of different segmentation models are provided in Figure 5. Semantic segmentation models encounter difficulties when segmenting closely located eggs. In certain instances, these models generate false positive predictions for eggs. While our proposed architecture has mitigated this issue to some extent, complete elimination remains elusive. Some areas in the predictions still exhibit noise or incomplete segmentation, arising from the partial segmentation of a portion of an egg.

Since the instance segmentation, and Watershed models provided a lower accuracy than expected, we also qualitatively compare those models in Figure 6. One notable

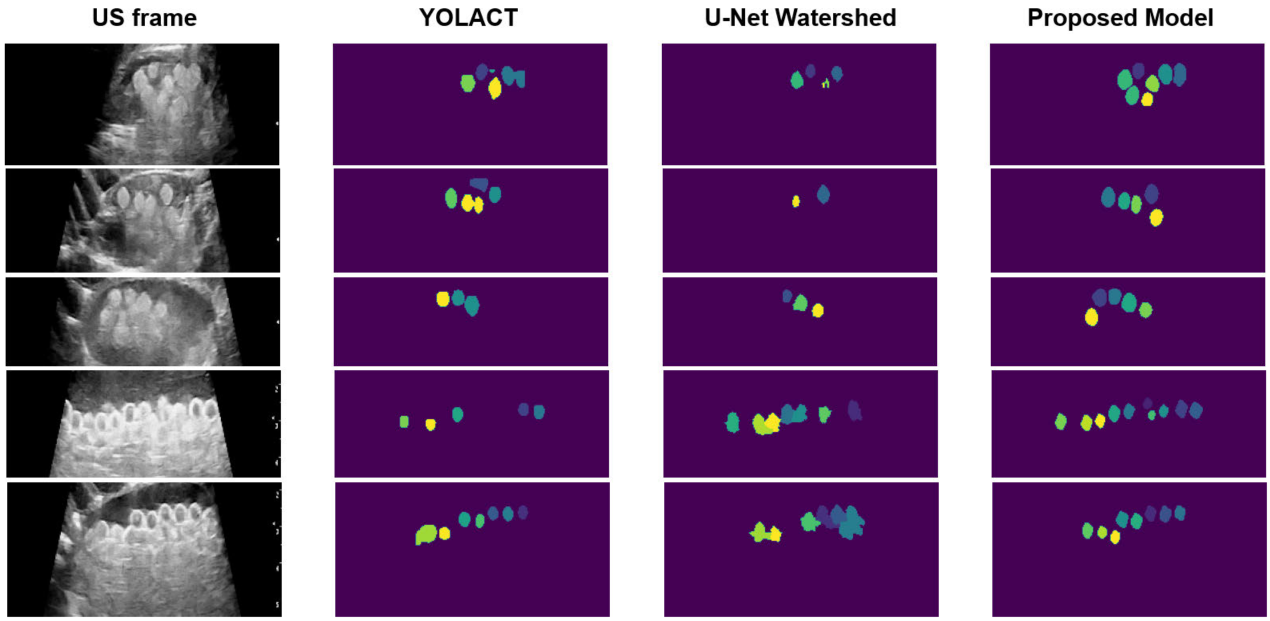

FIGURE 6. Comparison of instance segmentation, Watershed technique, and proposed model in salmon egg segmentation.

issue revolves around boundary detection in the Watershed technique, which exhibits considerable inconsistency. The use of Watershed for egg separation does not improve the segmentation results and it can not preserve the egg shape, causing over-segmentation and inconsistent Watershed line separation.

As for the instance segmentation, we observed instances of egg mask intersections in the YOLACT model, even though one of the primary motivations for incorporating instance segmentation into this application was to mitigate such problems. Another problem regarding instance segmentation was its tendency to generate smaller masks for individual eggs. This phenomenon also occurs in our proposed model, although it appears to be even more prevalent in the case of instance segmentation.

Our proposed model, with and without Attention, most efficiently addresses the problem of overlapping objects in the comparison. The utilization of Attention mechanisms in the proposed model enhances egg segmentation but still introduces some noise and occasional false positive predictions. The proposed model without Attention improves on these issues as can be seen in Figure 6. We, therefore, chose the model without Attention as our main model for our use cases of egg maturation state prediction and egg size measurement.

To evaluate the agreement between the average egg size in the manual measurement, estimated by operators using ultrasound data, and the proposed model, we are presenting the Bland–Altman analysis in Figure 7. While outliers are present, it's reassuring to observe that the majority of the test data comfortably aligns within the boundaries of the confidence intervals. The mean difference value hints at a relatively modest disparity between the manual measurement

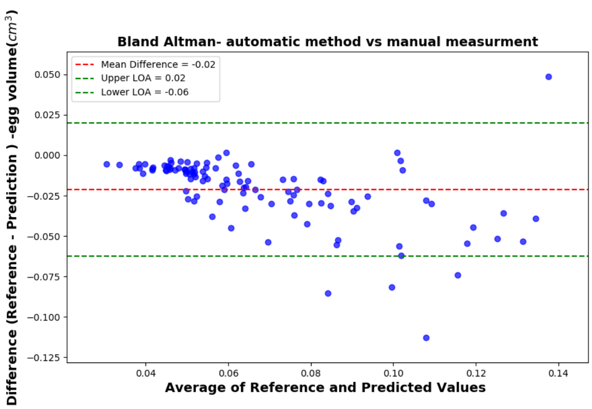

FIGURE 7. The egg size manual measurement (Reference) comparison to the model egg size measurement(Prediction).

and the deep learning model. Furthermore, the plot shows a trend toward overestimation and an increasing level of disagreement for larger eggs.

The proposed model showed a high potential for Atlantic salmon egg morphology assessment application and was tested in a live-field setting for initial verification as shown in Figure 8, providing in real time the average egg size, number of segmented objects, and egg maturation stage.

VI. DISCUSSION

The overall segmentation result shows that the proposed approach provides good accuracy in salmon egg segmentation and classification.

The Attention gates struggle with small egg segmentation and increase the classification errors in the image. This is known for small object segmentation because in the non-local Attention, vital details of small objects can be neglected or

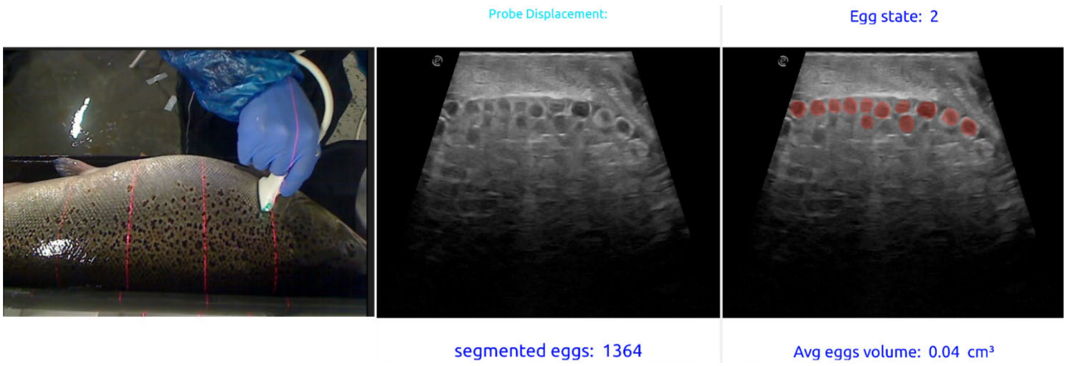

FIGURE 8. Real-time egg morphology application.

lost within abstract or coarse feature maps [28]. A potential solution to this issue is to use deep supervision Attention [29] or a multi-dimensional Attention [28]. However, employing deep supervised area-based Attention might not fully resolve the noise problem, given that the Attention is directed within the ovary. Eggs situated at lower depths in the ovary may experience substantial attenuation, making their distinction challenging. Another issue with using Attention in our proposed model is that balancing the segmentation and classification heads becomes particularly challenging due to the emphasis on segmentation-specific features in Attention, which may have an impact on classification performance. While conducting experiments with Attention, our main focus was on evaluating the efficacy of egg detection in the lower ovary parts rather than solely comparing the quantitative performance of two different architectures in our proposed model. Both versions of the proposed model have demonstrated satisfactory performance.

We were not able to improve our results using the YOLACT approach for instance segmentation. Upon an examination of the classification and segmentation results, it seems that the primary challenge lies in accurately classifying eggs. In the binary segmentation comparison, the YOLACT segmentation results are not comparable to the proposed model. While YOLACT generally provides satisfactory segmentation quality, there are instances where the masks for closely located eggs intersect. This issue could be attributed to several factors, with anchor boxes or proposals and Non-Maximum Suppression (NMS) standing out as potential causes for the intersections. In our YOLACT architecture, we employ ResNet50 for the FPN feature extraction. While there has not been such an accuracy drop in the proposed model by switching from convolution blocks to ResNet30's residual blocks for feature extraction, it is unlikely that feature extraction (FPN) is the root cause of the problem within YOLACT. A more plausible culprit is that the anchor boxes, lack of localization step, or Proposals may not be appropriately sized or spaced to handle closely packed objects effectively. YOLACT, like many other instance segmentation models, relies on Non-Maximum Suppression (NMS) to eliminate redundant bounding boxes. In situations where a bounding box mistakenly contains two or more eggs, it may be prioritized over other underlying bounding boxes, resulting in intersected mask objects. This inadequacy can lead to challenges in generating accurate masks.

Utilizing the Watershed technique, while computationally efficient, can lead to issues such as over-segmentation and imprecise separation of eggs. This occurs because Watershed lines don't conform smoothly to the actual contours of the eggs, making it challenging to achieve accurate foreground and background component separation.

Upon examining Figure 7, it becomes apparent that there is increased variability in the differences, particularly for larger eggs in the maturation phase. This variability suggests a potential inconsistency in the agreement between the proposed size estimation methods and manual measurements. The observed variation may be attributed to an overestimation of egg size by the prediction model, as evidenced by the mean difference. Furthermore, accurately annotating rounded objects such as eggs during the manual measurement process presents challenges, contributing to the observed dispersion. Annotators might lean towards underestimating egg boundaries due to the difficulty in precisely representing the contours of rounded compact shapes. This tendency results in fewer spline annotation points on eggs, which may not perfectly align with their actual shapes. Additionally, the proposed model tends to segment more eggs than manual measurement in the majority of instances. This leads to an average egg size calculated from a greater number of eggs, causing an overall size difference compared to manual measurements.

The model's performance on larger eggs could be improved with additional training data and a more comprehensive dataset. The presence of less frequent large eggs, contributing to the bias observed in the Bland-Altman analysis, prompts further investigation to discern whether the variation is genetically influenced or stems from other factors. This focused examination can provide valuable insights into the model's limitations and guide potential refinements for more accurate size estimations, especially in the context of larger eggs.

Further evaluation of the model, particularly through testing against ground truth values on a larger scale, can offer additional insights and enhance the overall assessment of the model's performance performance.

In summary, the real-time application yields promising initial results, allowing continuous monitoring of egg maturation stage and size throughout the harvesting period. Deploying the system for further training and testing in the field would contribute to further evaluation and verification. This could be a step toward automating the measurement of the maturation state in Atlantic salmon.

VII. CONCLUSION

This paper presents a deep learning application designed to estimate egg size and egg developmental stage in Atlantic salmon. The model segments closely situated individual eggs within the ovary and categorizes them into four distinct developmental stages. This achievement offers a valuable tool for assessing salmon egg maturation, enabling the study of egg size and maturity stage variation for continuous monitoring during harvesting. The application has the potential to replace manual measurements and palpation processes in fish farming facilities. The application has been tested in an operational environment with promising initial results. Future work will focus on further testing on a larger scale and improving the current method.

ETHICAL STATEMENT

This research has been carried out while adhering to standards outlined by the Norwegian aquaculture legislation by the professionals at MOWI/AquaGen.

ACKNOWLEDGMENT

The authors would like to thank the staff with AquaGen, MOWI, and InPhase Solutions for their support and also would like to thank Erik Smistad for his technical input to this manuscript. The authors used Grammarly and ChatGPT language models to improve the grammar and overall text quality.

REFERENCES

[1] I. Næve, S. A. Korsvoll, N. Santi, M. Medina, and A. Aunsmo, "The power of genetics: Past and future contribution of balanced genetic selection to sustainable growth and productivity of the Norwegian Atlantic salmon (*Salmo salar*) industry," *Aquaculture*, vol. 553, May 2022, Art. no. 738061.

[2] I. Naeve, M. Mommens, A. Arukwe, and E. Kjørsvik, "Ultrasound as a noninvasive tool for monitoring reproductive physiology in female Atlantic salmon (*Salmo salar*)," *Physiol. Rep.*, vol. 6, no. 9, May 2018, Art. no. e13640.

[3] I. Næve, M. Mommens, A. Arukwe, J. Virtanen, M. E. Hoque, and E. Kjørsvik, "Ultrasound as a noninvasive tool for monitoring reproductive physiology in male Atlantic salmon (*Salmo salar*)," *Physiological Rep.*, vol. 7, no. 13, Aug. 2019, Art. no. e14167.

[4] L. J. Belaid and W. Mourou, "Image segmentation: A watershed transformation algorithm," *Image Anal. Stereol.*, vol. 28, no. 2, pp. 93–102, 2009.

[5] A. Mangan and R. Whitaker, "Partitioning 3D surface meshes using watershed segmentation," *IEEE Trans. Vis. Comput. Graphics*, vol. 5, no. 4, pp. 308–321, Oct. 1999.

[6] L. Shafarenko, M. Petrou, and J. Kittler, "Automatic watershed segmentation of randomly textured color images," *IEEE Trans. Image Process.*, vol. 6, no. 11, pp. 1530–1544, Nov. 1997.

[7] Y.-L. Huang and D.-R. Chen, "Watershed segmentation for breast tumor in 2-D sonography," *Ultrasound Med. Biol.*, vol. 30, no. 5, pp. 625–632, May 2004.

[8] H. Zhang, Z. Tang, Y. Xie, X. Gao, and Q. Chen, "A watershed segmentation algorithm based on an optimal marker for bubble size measurement," *Measurement*, vol. 138, pp. 182–193, May 2019.

[9] V. Yu. Volkov, "Compact objects detection in noisy images," in *Proc. 11th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2022, pp. 1–4.

[10] J. Zhang, Z. Zhou, and J. Sun, "Center-aware instance segmentation for close objects in industrial 3-D point cloud scenes," *IEEE Trans. Indus. Informat.*, vol. 20, no. 2, pp. 2812–2821, Feb. 2024.

[11] X. Fei, X. Li, C. Shi, H. Ren, I. Mumtaz, J. Guo, Y. Wu, Y. Luo, J. Lv, and X. Wu, "Dual-feature fusion attention network for small object segmentation," *Comput. Biol. Med.*, vol. 160, Jun. 2023, Art. no. 106985.

[12] S. Sang, Y. Zhou, M. T. Islam, and L. Xing, "Small-object sensitive segmentation using across feature map attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 5, pp. 6289–6306, May 2023.

[13] L. Zhang, K. Zhang, and H. Pan, "SUNet++: A deep network with channel attention for small-scale object segmentation on 3D medical images," *Tsinghua Sci. Technol.*, vol. 28, no. 4, pp. 628–638, Aug. 2023.

[14] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101563.

[15] Y. Yari, I. Næve, A. Hammerdal, P. H. Bergtun, S.-E. Måsøy, M. M. Voormolen, and L. Lovstakken, "Automated measurement of ovary development in Atlantic salmon using deep learning," *Ultrasound Med. Biol.*, vol. 50, no. 3, pp. 364–373, Jan. 2024.

[16] E. Smistad, A. Østvik, and L. Løvstakken, "Annotation web—An open-source web-based annotation tool for ultrasound images," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2021, pp. 1–4.

[17] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, in Lecture Notes in Computer Science, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham, Switzerland: Springer, 2014, pp. 740–755.

[18] D.-H. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Proc. Workshop Challenges In Represent. Learn.*, vol. 3, no. 2, 2013, p. 896.

[19] S. Leclerc, E. Smistad, J. Pedrosa, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. R. Berg, P.-M. Jodoin, T. Grenier, C. Lartizien, J. D'hooge, L. Lovstakken, and O. Bernard, "Deep learning for segmentation using an open large-scale dataset in 2D echocardiography," *IEEE Trans. Med. Imag.*, vol. 38, no. 9, pp. 2198–2210, Sep. 2019.

[20] P. Liu, X. Qiu, and X. Huang, "Adversarial multi-task learning for text classification," 2017, *arXiv:1704.05742*.

[21] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.

[22] W. E. Higgins and E. J. Ojard, "Interactive morphological watershed analysis for 3D medical images," *Computerized Med. Imag. Graph.*, vol. 17, nos. 4–5, pp. 387–395, Jul. 1993.

[23] J. Wang, H. Lu, G. Eude, and Q. Liu, "A fast region merging algorithm for watershed segmentation," in *Proc. 7th Int. Conf. Signal Process. (ICSP)*, Aug. 2004, pp. 781–784.

[24] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time instance segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 9156–9165.

[25] E. Smistad, M. Bozorgi, and F. Lindseth, "FAST: Framework for heterogeneous medical image computing and visualization," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 11, pp. 1811–1822, Nov. 2015.

[26] E. Smistad, A. Østvik, and A. Pedersen, "High performance neural network inference, streaming, and visualization of medical images using FAST," *IEEE Access*, vol. 7, pp. 136310–136321, 2019.

[27] F. Lux and P. Matula, "Cell segmentation by combining marker-controlled watershed and deep learning," 2020, *arXiv:2004.01607*.

[28] B. Wang, L. Wang, J. Chen, Z. Xu, T. Lukasiewicz, and Z. Fu, ''W-Net: Dual supervised medical image segmentation model with multi-dimensional attention and cascade multi-scale convolution,'' 2020, *arXiv:2012.03674*.

[29] D. Bouget, A. Pedersen, S. A. M. Hosainey, O. Solheim, and I. Reinertsen, ''Meningioma segmentation in T1-weighted MRI leveraging global context and attention mechanisms,'' *Frontiers Radiol.*, vol. 1, Sep. 2021, Art. no. 711514.

YASIN YARI (Graduate Student Member, IEEE) received the M.S. degree in systems engineering with embedded systems from the University of South-Eastern Norway. He is currently pursuing the Ph.D. degree with Norwegian University of Science and Technology. His research interests include machine learning and computer vision, embedded real-time systems, and data science.

INGUN NÆVE received the Ph.D. degree in biology from Norwegian University of Science and Technology (NTNU). She is currently a Fish Biologist and the Head of Broodstock and Genetics with the Lerøy Seafood Group, one of the world's largest seafood companies. She developed non-invasive methods using ultrasound to monitor the reproduction of Atlantic salmon with NTNU. Her research was conducted in collaboration with AquaGen, a leading company in breeding and genetics of salmon and trout. She was a Researcher with AquaGen, where she applied her expertise in ultrasound technology to improve the welfare and performance of broodstock candidates. Her research interests include the development and optimization of broodstock management, especially in relation to sex sorting, maturation monitoring, and egg quality. She has published several scientific articles and articles on these topics. Her passion is working for a sustainable aquaculture industry and she enjoys working hands-on in production with her colleagues.

PER HELGE BERGTUN received the Veterinarian degree from Norwegian School of Veterinary Science. He has been working in fish health with MOWI, since 2005. He started in the Breeding Division, in 2016, as the Fish Health Manager with interest in fish health and reproduction biology.

ASLE HAMMERDAL received the joint M.S. degree in cybernetics engineering from Norwegian University of Science and Technology (NTNU) and the University of California Santa Barbara. He is currently a Senior Researcher with AquaGen AS.

SVEIN-ERIK MÅSØY (Member, IEEE) was born in Halden, Norway, in 1975. He received the M.Sc. degree in applied mechanics, thermodynamics, and fluid dynamics from Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1999, and the Ph.D. degree, in 2004. His Ph.D. dissertation titled ''Estimation and Correction of Aberration in Medical Ultrasound Imaging.'' He is currently an Associate Professor with the Department of Circulation and Medical Imaging, NTNU. His research interests include beamforming, image processing, machine learning, nonlinear acoustics, and non-destructive testing using ultrasound.

MARCO MARIEN VOORMOLEN was born in Utrecht, The Netherlands, in 1972. He received the M.Sc. degree in measurement and control engineering and in biomedical engineering from the University of Twente, The Netherlands, in 1998. Since 1999, he has been a Consultant on safety & control systems for the railway industry. In 2001, he joined the Department of Biomedical Engineering, Erasmus University Medical Center, Rotterdam, as a Scientific Researcher where, in addition, he successfully pursued his Ph.D. degree. His Ph.D. dissertation titled ''3D Harmonic Echocardiography,'' involved non-linear acoustic theory, 3D echocardiographic imaging techniques, and image processing. In 2008, he moved to Trondheim to work as a Post-Doctoral Researcher with Norwegian University of Science and Technology (NTNU), where he focused on transducer technology for high-frequency ultrasonic applications. From 2011 to 2013, he was employed with Aurotech Ultrasound, where he was involved in regulatory affairs and management of client portfolios. In 2013, he joined InPhase Solutions as a co-owner and currently holds the position of the CTO and the Chair of the Board.

LASSE LØVSTAKKEN (Member, IEEE) is currently a Professor with Norwegian University of Science and Technology (NTNU). He is also the Head of the Ultrasound Research Group, St. Olavs University Hospital. He has supervised 25 master's students and 14 Ph.D. students (eight graduated) as a main supervisor and 11 as a co-supervisor. He is also an Expert in ultrasound signal and image processing/analysis. His work on vector flow imaging has led to several innovations available in commercial clinical ultrasound scanners.

• • •