Irregularity detection in net pens exploiting Computer Vision ⋆

Christian Schellewald ** **Annette Stahl** *

* *Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), O. S, Bragstads plass 2D, 7491 Trondheim, Norway (e-mail: Annette.Stahl@ntnu.no).*
** *SINTEF Ocean AS, Brattørkaia 17C, 7010 Trondheim, Norway (e-mail: Christian.Schellewald@sintef.no).*

Abstract: Protecting the remaining wild salmon stock in Norway is of utmost importance and requires that farmed salmon cannot escape from aquaculture sites. As holes in net-cages are responsible for a large fraction of the escaped salmon the industry has to perform frequent inspections of the fish cage integrity. In this paper we propose an image processing and computer vision based attention mechanism towards a more automated fish-cage inspection. The presented algorithm allows to indicate areas in videos showing net-pen locations where potential holes are present. We show the effectivity of the approach on video-recordings of holes also in commercial fish-cages.

Keywords: Hole detection, Computer Vision, Morphological Filter, ROV, Fish cage, Net detection, irregularity detection

1. INTRODUCTION

Within modern offshore based aquaculture, using gravity net cages, net malfunction and escape of fish is a threat to wild fish populations. From 2010 to 2018, a total of 305 reported escape incidents with Atlantic salmon (Salmo salar) or rainbow trout (Oncorhynchus mykiss) were confirmed, involving in total 1.960.000 registered escapees (Føre and Thorvaldsen, 2021). To minimize escapes due to malfunctioning of net cages, inspections must be performed regularly to discover and repair damages as early as possible, minimizing the time period with a non-intact net structure. Today, net inspection is performed manually by divers or farming personnel monitoring video streams recorded by Remotely Operated Vehicles (ROVs) tethered to, and operated from, a surface vessel. The ROV is thereby remotely controlled to move in horizontal or vertical direction close to the net while streaming the video footage to a boat or operation center of the farming site. Typical net cages are circular with a radius of approximately 25m and a depth of up to 50m (Jensen et al., 2010). Using manual labor for inspecting such large net structures is tedious and expensive. Hence, the inspection frequency is limited by economic factors, and typically carried out on a monthly basis, or after any operation involving the net and weighting system (official requirement for Norwegian fish farming companies). Manual operation also makes the inspection quality highly operator dependent. Technology that allows for an inspection to be carried out by an autonomous underwater vehicle (AUV) was for example proposed in Rundtop and Frank (2016), Potyagaylo et al. (2015), Duda et al. (2015) and Schellewald et al. (2021) but full autonomy still requires some development efforts before it will be reached.

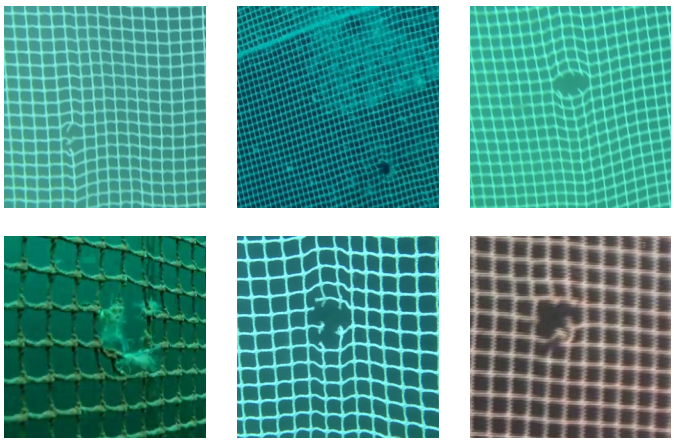

Fig. 1. Net damage examples. Detecting holes manually within the regular net-structure is a tedious and expensive task.

An AUV could be deployed from an underwater docking station or manually from a surface vessel, and hence reduce labour cost, allow for increased inspection frequency, avoid operator dependent inspection results and allow for a systematic documentation of the net cage integrity. Within our presented work the focus is on the development of a machine vision system for net damage detection. This is one of the key enabling technologies for the realization of an autonomous inspection vehicle. Moreover, such a system can be utilized as aid for the ROV-pilot indicating possible mesh breakages. Within salmon aquaculture, the main reasons for escapes are structural breakdowns due to environmental loads, operational causes like erroneous

⋆ This work was financed by the Research Council of Norway through the project SFI Exposed (237790) and the Centre of Excellence funding scheme, project number 223254, AMOS.

mounting, over straining or propeller/boat damage, abrasion between net and chain or ropes. Other damages come from predators and external causes like flotsam/wreckage or ice (Jensen et al., 2010). Depending on the cause, net damages have different shapes and sizes (cf. Figure 1). However, the most common type of failure is a single mesh breakage. Detecting and repairing them is important to avoid a subsequent enlargement of the hole. Larger holes require immediate attention to avoid potential escape of fish. Therefore, an important scenario towards an autonomous net damage detection system is the ability to detect small holes in a single net layer within a fish cage. The challenges in automatic processing (Schellewald et al., 2021) and analysis of underwater net cage structure are manifold (Duda et al., 2015); due to the turbidity of the water, caustics, reflections along with possible low light conditions, the video quality might be poor even with high quality cameras. Additionally, the water current and waves can cause spatial deformations in the net structure, fish can regularly occlude the net, and algae growth often covers the net structure to a certain degree leading to net irregularities (Madshaven et al., 2022).

1.1 Related Work

Automatic inspection operations of net cage structure integrity utilize different sensor modalities such as cameras or acoustic based sensors. The computer vision-based approaches by (Zhao et al., 2020; Paspalakis et al., 2020; Jovanovi et al., 2016) utilize image processing techniques like Otsu's thresholding method (Otsu, 1979) to generate a binary image and the Hough Transform for line detection (Illingworth and Kittler, 1988). The pixel count for each grid cell was determined, indicating a possible net structure damage if the count is largely above average. A modified version followed where line detection was used to calculate the nearest net structure pixel within a binary image mask. A similar approach combining both techniques were proposed by (Betancourt et al., 2020) to detect if net junctions are missing. The method was tested on real fish cage image footage with a detection accuracy of 79%. The team of (Zacheilas et al., 2021) proposed a FPGA-based system utilizing image enhancement and image processing methods. In contrary, this work focuses on the regularity of the net pattern without accounting for the patch size of the connected components.
Another category of methods are based on rejecting certain feature detections on dynamic objects, like fish or floating seaweed (Leonardi et al., 2020) as they complicate the process of reconstructing mesh structures such as knot points, and in certain situations resemble net holes. Image enhancing methods and marine growth segmentation were applied by (Qiu et al., 2020). Several underwater images were collected and labeled with pixel-wise annotations. Deep learning for net hole detection was applied by (Tao et al., 2018). YOLO (Redmon et al., 2016) was utilized for net structure and hole detection applied on images taken under controlled lab conditions. (Madshaven et al., 2022) utilized an ensemble of classical computer vision and image processing techniques for tracking and made use of neural networks for both segmentation of net structure and classification of scene content, as the determination of irregularity (fish, seaweed).

1.2 Contributions

In this paper a novel approach is proposed that targets the net irregularity detection, a specific task needed within a full fish-cage inspection processing chain. In particular, an attentional mechanism is presented that allows to mark areas where potential holes are present in an otherwise regular net-structure. The underlying multi-stage regularity check robustifies the selection process and declares different levels of potential net irregularities/damages. The proposed method works under realistic lighting condition and was tested by using real world aquaculture fish net recordings. The entire approach facilitates traditional computer vision and image processing methods, preventing time consuming and expensive training as performed when using modern machine learning paradigms.

1.3 Overview

In the following, the suggested approach is explained in detail by presenting in section 2 the main idea as well as the needed theoretical background knowledge. After that, in section 3, the algorithmic steps are explained, followed by the result-section (section 4) along with a discussion and the conclusion (section 5).

2. THEORETICAL BACKGROUND

The aim of the presented approach is to provide an attentional mechanism that allows to mark areas where potential holes are present in an otherwise regular net-structure. This section explains the main idea and briefly reviews some image processing techniques are exploited in the approach. Then the algorithm is explained in more detail.

2.1 Main idea

The main idea followed to automatically detect the presence of potential holes in the net-structure of a fish cage is based on the fact that large areas of the net pen show a regular net pattern. As a potential hole represents a deviation of this regular pattern at first the regular structure of the net is determined followed by a search for locations in the image where the regularity of the neighborhood is no longer present. This is done by testing if the locally observed regular pattern continues into the close neighborhood and keep track of locations where the regularity is broken. The approach copes with two common lighting

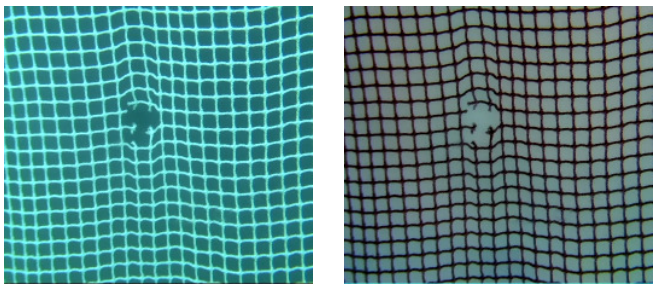

Fig. 2. Video frame of a hole in a net pen. The **left** image is the original where the net appears brighter than the background. The **right** image is the result of a morphological "black top-hat" operation which – beside reducing the effect of a potential brightness gradient – also inverts the image.

situations. In the first the net appears brighter than the

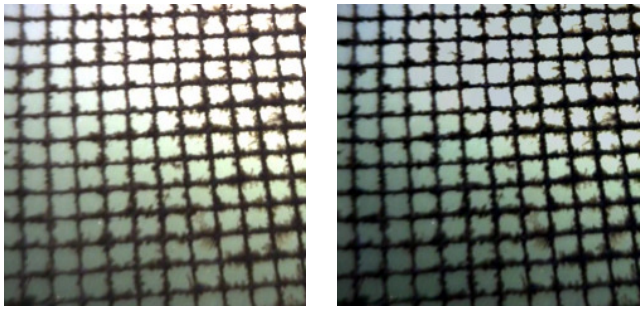

Fig. 3. The **left** image is an original image of a net which appears darker than the background. The **right** image is the result of a morphological white top-hat operation enhancing the black structure and reducing the brightness gradient present in the background.

background and in the second the net appears darker than the background (see left images in Figure 2 and 3). After the application of two related morphological operations reducing the "brightness gradient" in both, the resulting images on the right side show net structure which appears dark in front of a brighter background. The incorporated inversion in the case of a bright net provides a common basis to extract the net structure by finding points with a local brightness maximum and using these for representing the mesh-structure of the net. As the proposed approach exploits morphological operations (i.e. black- and white-top-hat) they are introduced in the next section.

2.2 Morphological image processing

Morphological operations (Soille, 2003) are non-linear operations and are highly useful in computer vision to analyze and process geometrical and other structures in images. Here they are used to unify the subsequent processing steps by representing the net as dark structure when the original image contains a bright image and also to improve the separability of the net-structure from the background by enhancing the dark net-area and reducing the "brightness gradient" (transition from a brighter to a darker area) that might be present in the background.

Dilation and erosion

The two fundamental mathematical morphology operations are dilation and erosion which we denote as $\times_d$ and $\times_e$, repectively. They act on a structuring element $b(x)$ which consists of a set of pixels defining a neighborhood and an origin. The origin is typically one pixel of the neighborhood, although generally it can lie outside of the structuring element. In our case we use a small circular disk with the origin in the center as structuring element.
The *dilation* of image f by structuring element b is denoted as

$$f \times_d b. \tag{1}$$

Performing this operation the structuring element is swept over the image (like a filter-mask) and the new gray-value for each pixel is defined as the maximum of the pixels covered by the structuring element. Therefore, it extends bright structures in the image.
The *erosion* of image f by structuring element b is denoted as

$$f \times_e b. \tag{2}$$

Again, to perform this operation the structuring element is swept over the image but the new gray-value of each pixel is defined as the minimum of the pixels covered by the structuring element. This operation darkens the image. Definitions for basic morphological operations can be found in (Soille, 2003).

Closing and opening

Closing (denoted by $f \bullet b$) and opening (denoted by $f \circ b$) are two operations composed of erosion and dilation. The *closing* of an image f by a structuring element $b(x)$ is a dilation $\times_d$ followed by an erosion $\times_e$:

$$f \bullet b = (f \times_d b) \times_e b. \tag{3}$$

The *opening* of an image f by a structuring element $b(x)$ is an erosion followed by a dilatation:

$$f \circ b = (f \times_e b) \times_d b. \tag{4}$$

Black top-hat and white top-hat

In the case where the net is brighter than the background the black top-hat transform (sometimes referred to as bottom-hat transform) will be applied to invert the image (resulting in net-structure that is represented by dark pixels) and to reduce a potential present brightness gradient (cf. Figure 2).
The *black top-hat* transform is defined as the difference between the closing and the input image:

$$T_b(f) = (f \bullet b) - f. \tag{5}$$

The black top-hat returns an inverted image where structures like the net which are "smaller" than the structuring element become more clearly visible as dark net structure. In the case that the background detected is already brighter than the net structure a white top-hat transformation will be performed.
The *white top-hat* transform is defined as the difference between the input image and its opening by some structuring element:

$$T_w(f) = f - (f \circ b). \tag{6}$$

The white top-hat transform returns an image enhancing the dark net structure and reducing a potential present brightness gradient (cf. Figure 3).

3. APPROACH

In this section the single steps of the Algorithm are explained. This starts with the pre-processing steps followed by the determination of the regular net-structure to the determination of irregularities representing potential holes.

3.1 Algorithmic steps

De-interlacing: Videos from ROVs that are used for inspecting the fish cage are often transmitted as interlaced signal containing the information of two video frames captured consecutively. An example illustrating the artifacts resulting from interlacing is depicted in the left image of Figure 4. Using the same bandwidth interlacing allows to double the frame-rate at the cost of a reduced y-resolution (each second row belongs to the same time). For image processing the interlaced image has to be de-interlaced to obtain time consistent images. In cases where the video frames were interlaced we performed the de-interlacing by substituting every second row with the interpolated color-values from the neighboring rows (cf. right image of Figure 4).

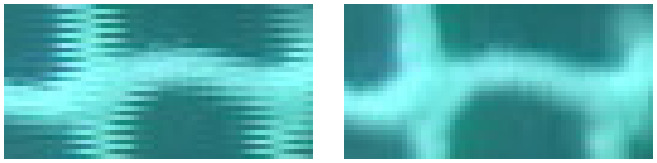

Fig. 4. De-interlacing. **Left:** An interlaced image. **Right:** De-interlaced version of the image.

Distinguishing dark and bright net: For distinguishing between the two illumination conditions a adaptive thresholding along with determining the size distribution of connected components was exploited. If many small localized dark/black segments are present and only a few wider spread bright/white connected components indicate the presents of a bright net (compare figure 5). The illumination case (bright/dark) determines which Morphological filter is used in the next step.

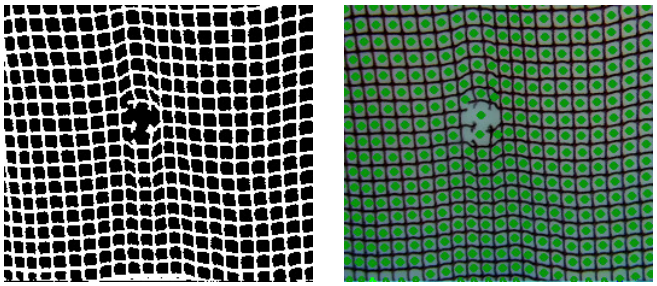

Fig. 5. **Left:** Binary segmentation of the net from Figure 2. **Right:**. Confirmed net structure.

Morphological pre-processing: As net videos taken under real world conditions often show non-uniform illumination we apply the previous discussed morphological filters to the input image with the subsequent aim to robustly determine the center location of the net masks (containing the background) which are typically arranged in a regular fashion. In case of a bright net the black top-hat transform is used to reduce the effect of a potential brightness gradient and also inverts the image such that the net-structure appears dark and the background bright. In the case where the net already appears dark we employ the white top-hat transform. Both filters improve the separability of the net structure from the background (see right images of Figure 2 and Figure 3).

Non-maximum suppression: This step allows to find the bright center locations of the net cells by searching for all local brightness maxima in a smoothed version of the image. We do so by employing a non-maximum suppression (Neubeck and Van Gool, 2006) on the image. The non-maxima-suppression on gray-value images finds all pixels that belong to a local maximum including the 'plateaus' with a value equal to the local maximum. We mark the centroids of the maxima/plateaus as local maximum. This helps to determine the location of the net cells, particular in real world recordings, in a efficient and robust way. Figure 6 shows the results of the non-maxima-suppression when applied to the right image of Figure 2. The subsequently extracted grid-points of the net consist of the centroids of the found maxima/plateaus.
We note that the non-maximum suppression for extracting the local maxima can be implemented efficiently using a few basic morphological operations along with a smoothing

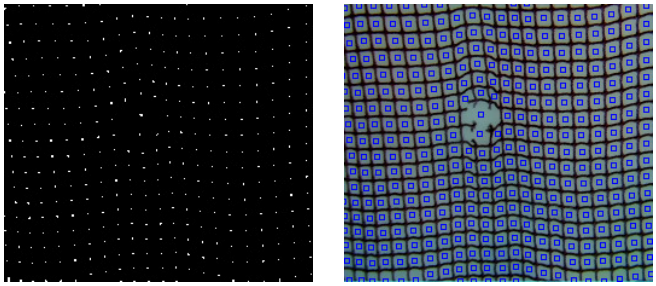

Fig. 6. **Left:** Binary image showing the pixels belonging to the maximum plateaus extracted by a non-maxima-suppression. **Right:** Extracted grid-points of the net are the centroids of the maxima/plateaus.

and comparison operation: First, the image is converted to a gray-value image and eroded using a disk shaped structuring element and subsequently smoothed with a Gaussian kernel. The resulting image is then subtracted from a dilated version of this image. Resulting zero-valued image regions (pixels) represent the desired local maxima positions.

The next step aims to determine which of the found local maxima agree to lie on a local regular grid and simultaneous determine where the grid-structure is not present.

k-nearest neighbor (kNN) testing: Starting from the obtained local maxima positions a regularity check is performed exploiting a k-nearest neighbor (kNN) test verifying the regular grid structure of the net. The kNN

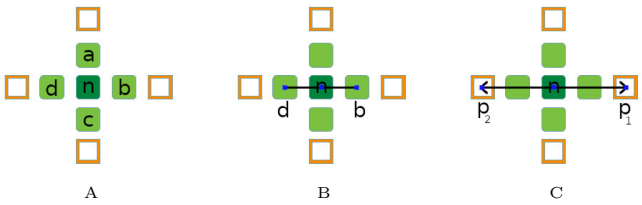

Fig. 7. Regularity check of the maxima nodes. A) The 4 nearest neighbours (a, b, c, d) of the currently considered node n are determined. B) Opposed nodes (here d and b) are tested to lie approximately in line with node n. C) The vector between d and b is then used to postulate expected grid points p_1 and p_2. The same is done for the opposed nodes a and c that lie in the other direction (not illustrated).

search can be efficiently realized by arranging the maxima nodes into a kd-tree. Then for each node we find the 4 nearest neighbours. We illustrate this in Figure 7 for a single node n. By sorting these 4 neighbours (a, b, c, d) in clockwise order we can easily test if two opposed nodes (illustrated in horizontal direction d and b) and the center node n lie on a line by testing if the angles of the opposed points are approximately 180 degrees plus a tolerance (of about 23 degrees) apart from each other, in order to compensate for small deformations of the net. The vector between d and b is then used to postulate the positions of the nodes p_1 and p_2 that should be present in both directions in a regular net. If a local maximum is close by to the postulated node position we mark the found node as being "re-confirmed". If the postulated node is not present a break in the regular pattern might be present and the position gets marked as "non-confirmed". A similar

test is performed in the other direction defined by the nodes a and c. We mark the center nodes that build a regular cross with all their 4 neighbours as "confirmed". Regular net-grid structures (i.e. "confirmed" and "re-confirmed" nodes) are colored in green. The locations of non-reconfirmed points are checked once more. If the binary segmentation (into background and net-structure) indicates that net-structure is present at the location of a postulated node it gets the status of being irregular-net and is colored orange. If background is present at the postulated node location it gets the status of a confirmed-irregularity (colored blue). Locations where two or more irregularities are present in the close neighbourhood are marked as hole/irregularity and colored in red.

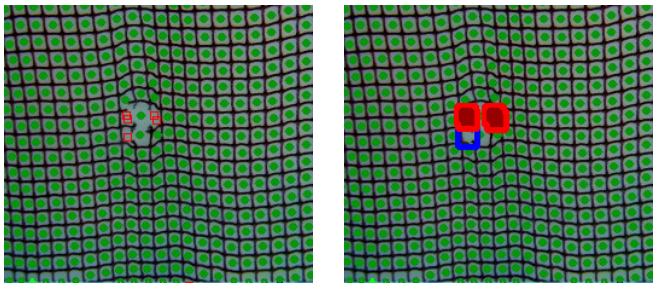

Fig. 8. k-nearest neighbor check: Net damage locations shown in red, irregular net structure shown in blue and regular net structures shown in green. **Left:** Regular net structure with locations of missing regularity. **Right:** Locations labelled as damaged net represent potential holes.

4. RESULTS

The proposed method was implemented in C++ exploiting the openCV library (Itseez, 2020) and the "Fast Library for Approximate Nearest Neighbors" FLANN (Muja and Lowe, 2009) allowing the algorithm to perform in real time (Videos with 512x512 pixels are typically processed with $>$ 20 fps using not highly optimized code). Results for videos from field experiments are shown in Figure 9. The proposed approach is able to detect the regular net structure and properly determines present irregularities. We wish to emphasize that the developed approach targets to solve a specific task needed within a full fish-cage inspection chain and accordingly has some limitations.

4.1 Limitations

In its current form, our algorithm detects all deviations from a regular net-structure. This can, for example, also be patches of seaweed from bio-fouling or other objects that directly occlude a part of the net (See Figure 10). Therefore, a subsequent object detection will be implemented in the future to further separate potential holes from non-holes. In addition, the algorithm currently performs a frame by frame analysis ignoring any time information. However, the tracking of potential holes over time will increase the robustness by helping to remove spurious detected irregularities and find holes/irregularities which are consistent over time. Another assumption of the approach is that the net is observed parallel to the net ($\pm$25 degrees) and that the size of the net-cells do not vary too much within one image frame due to the perspective projection. The second image row in Figure 10 indicates the challenges arising from a net tilted relative to the camera. Knowing the net orientation, a perspective correction could help to cope with that.

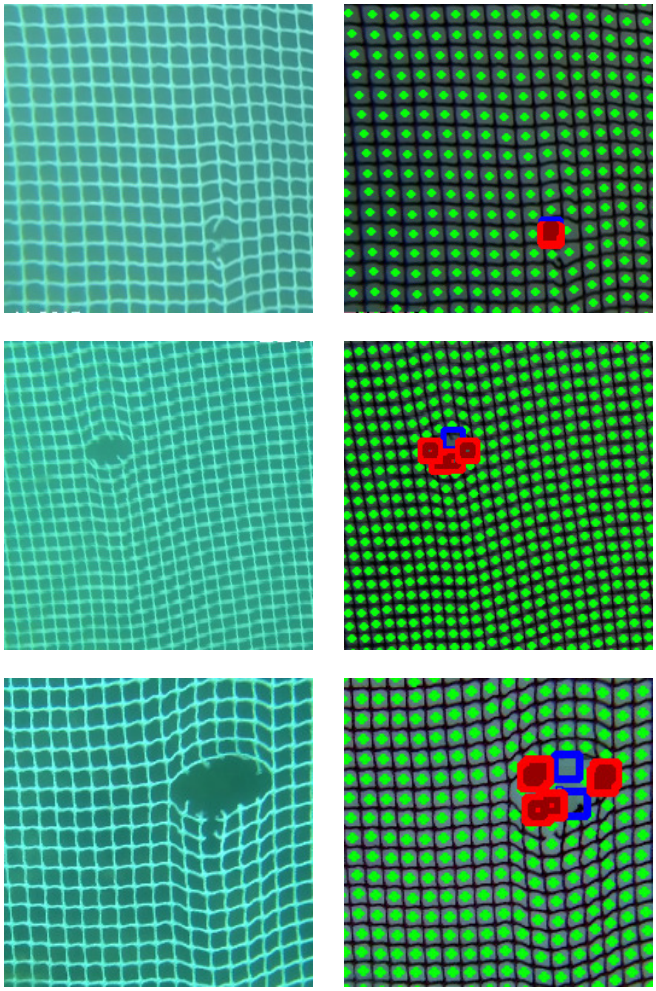

Fig. 9. Examples of net analysis results from videos of an ROV filming a net with different holes. Detected irregularities are marked.

5. CONCLUSION

In this paper, we presented an approach which is capable to detect irregularities in inspection videos from net-pens of salmon fish-cages. It allows for a more automated fish-cage inspection and is intended to be used as a module in underwater inspection vehicles that can traverse the net pen in an autonomous manner. The presented computer vision-based can serve as attention mechanism to indicate locations in videos showing net-pen where potential holes are present. The method is based on a two-stage regularity check where at first neighboring points of the center point are checked and in a second stage the connection vectors between two aligned grid points and the center node are used to postulate expected outer grid points. An additional consistency check leads to a more robust selection process for labeling net irregularities/holes. We showed the effectivity of the approach on video-recordings of holes also in commercial fish-cages.

ACKNOWLEDGEMENTS

We would also like to thank SINTEF ACE for enabling us to perform the trials using a ROV for recording the net pen videos from the boat Torra.

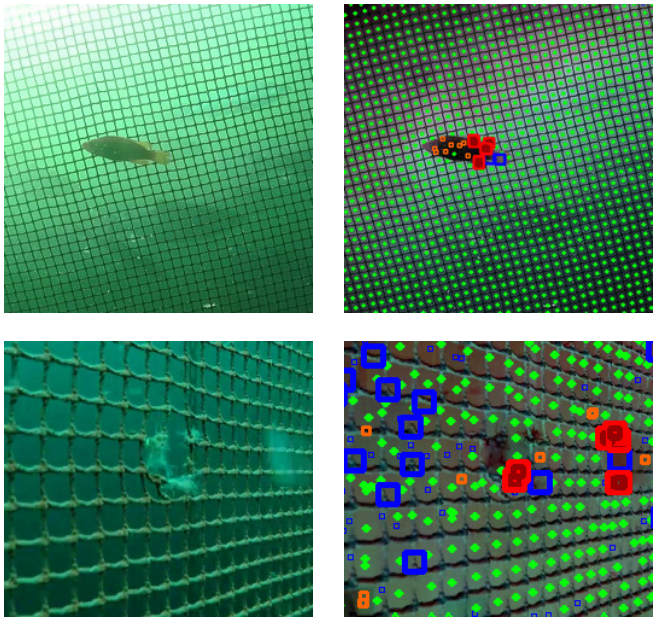

Fig. 10. Limitations: **Top row:** Objects occluding the net are detected as irregularities but represent no hole. **Bottom row:** The sizes of the net cells need to be similar which is violated in projections of a net tilted largely relative to the camera.

REFERENCES

Betancourt, J., Coral, W., and Colorado, J. (2020). An integrated ROV solution for underwater net-cage inspection in fish farms using computer vision. *SN Applied Sciences*, 2(12), 1–15.

Duda, A., Schwendner, J., Stahl, A., and Rundtop, P. (2015). Visual pose estimation for autonomous inspection of fish pens. In *OCEANS 2015 - Genova*, 1–6. doi: 10.1109/OCEANS-Genova.2015.7271392.

Føre, H.M. and Thorvaldsen, T. (2021). Causal analysis of escape of Atlantic salmon and rainbow trout from norwegian fish farms during 2010 - 2018. *Aquaculture*, 532, 736002. doi: https://doi.org/10.1016/j.aquaculture.2020.736002.

Illingworth, J. and Kittler, J. (1988). A survey of the Hough transform. *Computer vision, graphics, and image processing*, 44(1), 87–116.

Itseez (2020). Open source Computer Vision Library. `https://github.com/itseez/opencv`.

Jensen, Ø., Dempster, T., Thorstad, E.B., Uglem, I., and Fredheim, A. (2010). Escapes of fishes from Norwegian sea-cage aquaculture: causes, consequences and prevention. *Aquaculture Environment Interactions*, 1(1), 71–83.

Jovanovi, V., Risojevi, V., Babi, Z., Svendsen, E., and Stahl, A. (2016). Splash detection in surveillance videos of offshore fish production plants. In *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 1–4. doi:10.1109/IWSSIP.2016.7502706.

Leonardi, M., Fiori, L., and Stahl, A. (2020). Deep learning based keypoint rejection system for underwater visual ego-motion estimation. *IFAC-PapersOnLine*, 53(2), 9471–9477. doi: https://doi.org/10.1016/j.ifacol.2020.12.2420. 21th IFAC World Congress.

Madshaven, A., Schellewald, C., and Stahl, A. (2022). Hole detection in aquaculture net cages from video footage. In W. Osten, D. Nikolaev, and J. Zhou (eds.), *Fourteenth International Conference on Machine Vision (ICMV 2021)*, volume 12084, 258 – 267. International Society for Optics and Photonics, SPIE. doi: 10.1117/12.2622681.

Muja, M. and Lowe, D.G. (2009). Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *VISAPP (1)*, 2(331-340), 2.

Neubeck, A. and Van Gool, L. (2006). Efficient Non-Maximum Suppression. In *Proceedings of the 18th International Conference on Pattern Recognition - Volume 03*, ICPR '06, 850–855. IEEE Computer Society, Washington, DC, USA. doi:10.1109/ICPR.2006.479.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. doi: 10.1109/TSMC.1979.4310076.

Paspalakis, S., Moirogiorgou, K., Papandroulakis, N., Giakos, G., and Zervakis, M. (2020). Automated fish cage net inspection using image processing techniques. *IET Image Processing*, 14(10), 2028–2034. doi: https://doi.org/10.1049/iet-ipr.2019.1667.

Potyagaylo, S., Constantinou, C.C., Georgiades, G., and Loizou, S.G. (2015). Asynchronous UKF-based localization of an underwater robotic vehicle for aquaculture inspection operations. In *OCEANS 2015-MTS/IEEE Washington*, 1–6. IEEE.

Qiu, W., Pakrashi, V., and Ghosh, B. (2020). Fishing Net Health State Estimation Using Underwater Imaging. *Journal of Marine Science and Engineering*, 8(9).

Redmon, J., Divvala, S., Girshick, R., and 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)*.

Rundtop, P. and Frank, K. (2016). Experimental evaluation of hydroacoustic instruments for ROV navigation along aquaculture net pens. *Aquacultural Engineering*, 74, 143–156.

Schellewald, C., Stahl, A., and Kelasidi, E. (2021). Vision-based pose estimation for autonomous operations in aquacultural fish farms. *IFAC-PapersOnLine*, 54(16), 438–443. doi: https://doi.org/10.1016/j.ifacol.2021.10.128. 13th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2021.

Soille, P. (2003). *Morphological Image Analysis; Principles and Applications*. Springer-Verlag Berlin.

Tao, Q., Huang, K., Qin, C., Guo, B., Lam, R., and Zhang, F. (2018). Omnidirectional surface vehicle for fish cage inspection. In *OCEANS 2018 MTS/IEEE Charleston*, 1–6. doi:10.1109/OCEANS.2018.8604674.

Zacheilas, T., Moirogiorgou, K., Papandroulakis, N., Sotiriades, E., Zervakis, M., and Dollas, A. (2021). An FPGA-Based System for Video Processing to Detect Holes in Aquaculture Nets. In *2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE)*, 1–6. doi:10.1109/BIBE52308.2021.9635351.

Zhao, Y.P., Niu, L.J., Du, H., and Bi, C.W. (2020). An adaptive method of damage detection for fishing nets based on image processing technology. *Aquacultural Engineering*, 90, 102071.