Fuzzy Logic Approximation and Deep Learning Neural Network for Fish Concentration Maps

J. Mäkiö
University of Applied Sciences
Emden/Leer Faculty of Technology
Emden, Germany
juho.maekioe@hs-emden-leer.de

D. Glukhov
Computer systems and networks department
Polotsk State University
Novopolotsk, Republic of Belarus
d.gluhov@psu.by

R. Bohush
Computer systems and networks department
Polotsk State University
Novopolotsk, Republic of Belarus
r.bogush@psu.by

T. Hlukhava
Computer systems and networks department
Polotsk State University
Novopolotsk, Republic of Belarus
t.gluhova@psu.by

I. Zakharava
Computer systems and networks department
Polotsk State University
Novopolotsk, Republic of Belarus
i.zakharova@psu.by

Abstract— **This paper proposes an algorithm to obtain topographic maps of lakes, maps of fish concentration and a map of predator location based on the results of an intelligent sonar data processing. The algorithm is based on the following steps: input frame separation into overlapping blocks, blocks-processing using convolutional neural networks (CNN) YOLO v2, and merging extracted bounding boxes around one object. To construct maps of the distribution of features along the lake, we propose a novel method for constructing the approximation of GPS- referenced CNN results based on the original implementation of fuzzy logic.**

Keywords— sonar data; fish concentration; maps of lakes; fuzzy logic; convolutional neural networks

I. INTRODUCTION

Modern tools for detecting underwater objects with application of ultrasound (sonars) have become widespread in solving various applied problems. A highly specialized class of sonars designed to study the relief of the lake bottom and search for fish are called echo sounders. Such devices use ultrasonic waves of different ranges to construct a topographic map of the bottom. Due to the redundancy of information collected by echo sounders, the use of artificial intelligence technologies allows us to formalize and to use the experience and knowledge of experts in the analysis of echolocation data. The deployment of the best tool for image recognition based on deep learning neural networks allows us to talk about the use of the echo sounder for solving new applied problems, such as tourist, ecological, nature protection, search tasks.

In [1] authors propose sliding window filters with contour detection to extract low-level features and fishes contours on echo images. This approach cannot adapt to various shapes of fish-schools and bottom artifacts, because filters kernels were not specialized for complicated forms.

In [2] presented approach which use sliding window filtering to extract objects from echo image. At first a median filter was used for noise removing, after a low pass filter with adaptive threshold was used to separate tracks with fishes from the background noise level, finally a perimeter filter was used to remove small regions with echo pulses from stochastic noise and bottom-structure. Described method can give false negative results in case complicated forms school of fishes.

Next algorithm, presented in [3] proposes CNN approach to extract information about fish localization on echo image. This algorithm uses the sonar images of moving agent obtained by forward-looking sonar. Authors used CNN Yolo to binary classification. But in case high-resolution images and small-sized fishes this algorithm can give a lot of false negative results. Also, school of fishes can be missed, because it was not taken into account when learning CNN.

Traditionally, to surface a topology map building by a number of discrete measures, different kinds of interpolation methods are used. Geostationary estimation methods [4], like kriging, have large computational costs, but they can achieve optimal interpolation. In the case of echo data processing, it is important to take into account that the data are limited and frequently and the proper evaluation of static characteristics by these values is hard. GPS data updates rarely than the echo data that is why georefrencing is performing for group of echo sounding points. These groups are not equable located in water body and that is why we perform fuzzy logic for calculations.

This paper proposes a novel approach to generate fish concentration maps based on sonar data using CNN and that can adapt to different environment conditions. The proposed approach to detect fishes or other objects on sonar images is based on the following steps: 1) separation of the input image into overlapping blocks; 2) blocks-processing using CNN YOLO v2, and 3) merging extracted bounding boxes around one object. After fish detection, we propose a fuzzy logic based fish concentrations map building.

II. FISH DETECTION USING CNN

Deep machine learning systems provide perfect performance for object detection and classification challenges [5]. Object detection systems have to dedicate following contributions: accuracy, precise extraction of regions of interest (RoIs) on images, and their classification with minimal deviation and speed [6]. Typical image processing systems (optical character recognition system, fire detection video systems and others) usually include the following steps: preprocessing, features

extraction, classification, and context processing [7, 8]. Machine learning systems simulating the human brain, can solve detection and classification problem as good as or even better than the human brain. At the same time, machine learning systems are faster in problem solving than the human brain. Currently, CNNs are increasingly used for image processing in various practical areas. Unlike traditional networks, CNNs provide a reduced number of extracting parameters and as an alternative of whole image processing and can process only extracted feature map, which takes into account the image topology and is stable to affine transformation.

We analyzed famous neural network architectures, like AlexNet[8], Faster R-CNN [10], CoogleNet[11], ResNet[12] and etc. These networks process whole image as the feature map. This approach can make calculation faster, than the whole image processing. As stated in [13], YOLO is the fastest object detection system that works better than Faster R-CNN. In [14] the authors propose YOLO v2, YOLO9000, proposing modifications of YOLO. Better segmentation and classification was achieved by: 1) batch normalization from [15]; 2) high resolution classifier; 3) convolutional with anchor boxes; 3) use of k-means; 4) direct location prediction; 5) RPN usage; 6) fine-grained features; 7) multi-scale training; and 8) novel classification model Darknet-19.

CNN Darknet-19 has 19 convolutional and five max-pooling layers. It can distinguish 9000 classes. At training time, instead of fixing the input image size, the network was changed in every few iterations. After ten batches YOLO v2 randomly chooses a new image dimension size. Since this model down-sampled by a factor of 32, was pulled from the following multiples of 32: $\{320, 352, \dots, 608\}$. Input image resolution was resized to that dimension and continues training.

Since the sonar moves during the scanning of the lake along a complex trajectory with an alternating speed, it is necessary to perform the procedure of echogram normalizing. For this purpose, an algorithm to convert the echogram to metric coordinates along the length of the sonar track was developed. Due to the corresponding stretching/compression of the echogram, all objects of the acoustic echo may be represented on a single scale (fig.1).
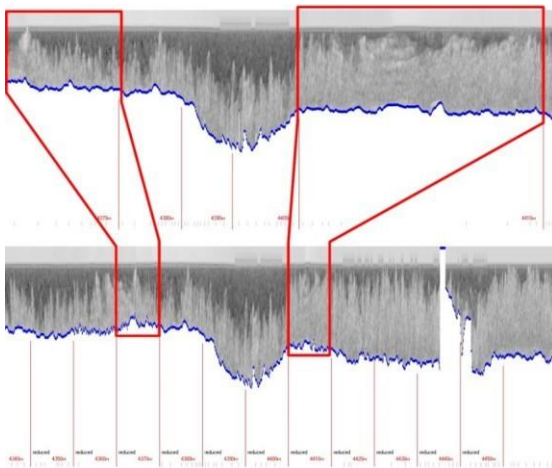

Fig. 1.   Example for echogram normalization process

The input images scaled before CNN processing. This means that tiny objects (fishes) can be missed. To solve this problem, we decided to process patches of echo image for precisely tiny

object detection and subsequently concatenate the output results with performing post processing actions.

We propose an effective algorithm for fish detection on sonar images based on the following steps: input frame separating into overlapping blocks, blocks processing using CNN YOLO v2, merging extracted bounding boxes around one object.

- Input image I with sizes $H \times W$ is divided into overlapping blocks $C_{i,j}$ with sizes $ch \times cw$, $i = \overline{0, H/ch-1}$, $j = \overline{0, W/cw-1}$ Overlap size can vary by input frame resolution and percentage ratio of minimal objects sizes.

- Each block goes to CNN YOLO v2 [15] in which network predict objects are localized by using sequences of convolutional filters. YOLO v2 uses convolution with anchor boxes, like in Faster R-CNN and run k-means (k=5) clustering for getting good priors for predicted objects. After YOLO v2 processing, we have bounding boxes in every block $C_{i,j}$ that are presented as top left corner coordinates $B_{i,j}(x_1, y_1)$, bottom right corner coordinates $B_{i,j}(x_2, y_2)$, object classification, and probability value.

- In the next step - blocks post-processing - the neighbor RoIs, which have combined overlapped region located closer than 10% from blocks edge, are searched. If these blocks are found, we calculate *IoU* (Intersection over Union) which describes two regions overlapping:

$$IOU = \frac{B_1 I B_2}{B_1 U B_2}, \qquad (1)$$

in which B_1 and B_2 are regions areas.

III.   MAPS BUILDING BASED ON FUZZY LOGIC

The idea of "fuzzy logic approximator" application to bottom topographical map building or for building feature map is based on following analogy: The assembly of separated depth measuring may be presented as knowledge-based systems including information about body water features and structure. The echo detection is described with formal logic as:

IF coordinates X, Y **AND** time t **THEN** depth D, water temperature T and other parameters

In this work the universal adaptive approximation is presented as fuzzy logic specific realization mathematical tools evolved in [16-19].

If distance between two points measuring operator is L(p,pi) then, as analogy of production system, water body fragment knowledge proliferates on neighbor fragments in accordance unimodal function has maximum value in specified point. We propose membership function in distance from nodal approximation point pi to point p as:

$$\varphi(p, p_i) = \frac{1}{L(p, p_i)^n} \left| L(p, p_i) < L_{\lim it} \right., \qquad (2)$$

In case application defuzzification by COG (Center Of Gravity) method, desired feature value for unknown point *p* can be calculated as:

$$p.z = \frac{\sum_i \varphi(p, p_i)\, p_i \cdot z}{\sum_i \varphi_i}, \qquad (3)$$

In addition, we present defuzzification methods modifications which overcome neighbor point's influence with help of space discretization. The replacement of rules group comes into discretization interval on one rule with maximum membership function in point *p* influence.

Fuzzy logic approximation, in contrast to traditional approximation methods, can take into account several predicates and build complex conditions.

For example, we can formulate approximation condition which allows not only depth but also information of structure of the lake bottom for correction abrupt depth change which may arise by bottom objects (artifacts).

IF coordinates X, Y **AND** bottom structure without, **THEN** depth D, water temperature T, and other parameters.

In this case predicates value needs to be normalized and we propose the following membership function modification:

$$\varphi(p, p_i) = \left(1 - \frac{L(p, p_i)}{L_{\lim it}}\right)^n \bigg| L(p, p_i) < L_{\lim it}, \qquad (4)$$

To define depth irregularity for point p as R(p), we normalize this value and estimate "bottom without artifacts" model as:

$$\varphi_R(p) = \left(1 - \frac{R(p) - R_{min}}{R_{max} - R_{min}}\right)^m, \qquad (5)$$

then
$$p.z = \frac{\int max\left(\min\left(\varphi(p, p_i), \varphi_R(p)\right)\right) z\, dz}{\int max\left(\min\left(\varphi(p, p_i), \varphi_R(p)\right)\right) dz} \qquad (6)$$

In conclusion, we get a classical logical output minimax representation about unknown parameter value and bottom approximation with low artifact influence.

IV. EXPERIMENTAL RESULTS

For Yolov2 we build our own training set including about 2000 objects. We selected ground truth bounding boxes around RoIs manually using VOTT (Visual Object Tagging Tool) software [18]. VOTT can make ground truth coordinates and convert them into Yolo format. Using this program, we additionally created annotations files. We predicted five classes of objects: "fish", "grass", "school of fish", "predator", "bottom fish". Figure 2 depicts ground truth boxes in VOTT.

a)   b)   c)   d)   e)

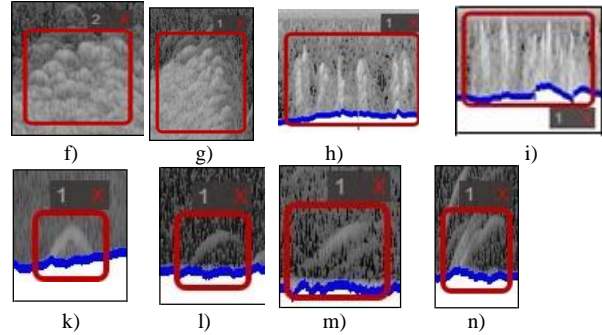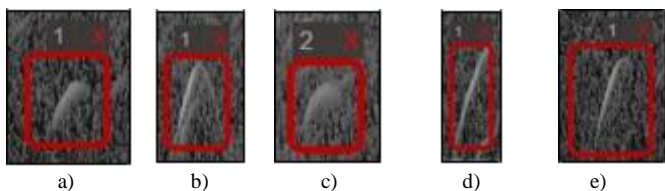f)   g)   h)   i)
k)   l)   m)   n)

Fig. 2. Ground truth bounding boxes in VOTT: a,b,c)"fish"; d,e) "predator"; f,g)"school of fish"; h,i)"grass"; k,l,m,n)"predator"

All images were taken from the river Western Dvina and lakes in the Republic of Belarus with a maximum depth for river of 12 meters and for lakes of 38 meters. Double-beam (200 kHz and 450 kHz) echo sonar Lowrance HOOK 4 was used. Figure 3 depicts the resulting classification after YOLOv2 processing.
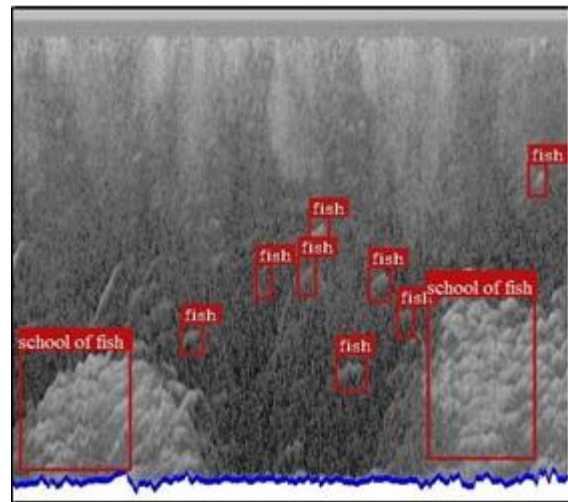

Fig. 3. Fish detection and classification results

Presented algorithm has an accuracy of 70.1% and a low percentage of false positive results in case of fish presence. However, our approach, as shown in Figure 4, cannot properly distinguish classes "grass" and "school of fish", especially in case similar shapes.

Depth map approximation is presented in Figure 5. Approximation is constructed using the radial membership functions. Therefore, the accuracy of the approximation increases with increasing number and density of approximation nodes. As the number of approximation nodes increases, the accuracy of approximation increases. In this case it is important that the sonar passes over all the most complex and characteristic sections of the bottom.
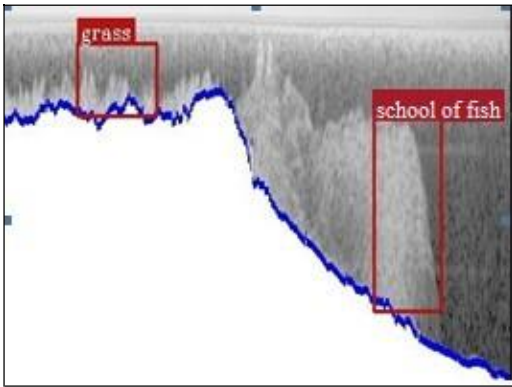
Fig. 4. Example for incorrect object classification

A special type of track is the lake contour, imported in KMZ format from known GIS systems, such as Google Earth. The contour is a track with zero-depth points (Figure 5a). Similar contours are used to simulate contours of islands. It is possible to delete incorrect points of the sonar tracks, as well as some points of the contours for modeling the open contour of the river bed. The same approach is used to construct the approximation by other features (Figure 5b).
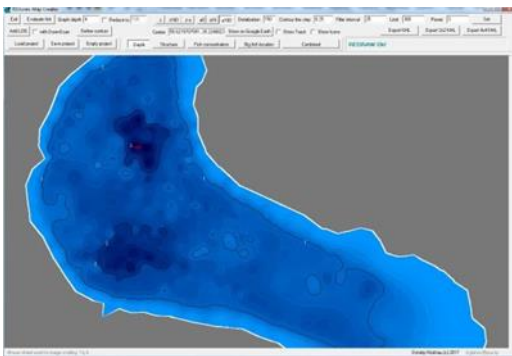
a)
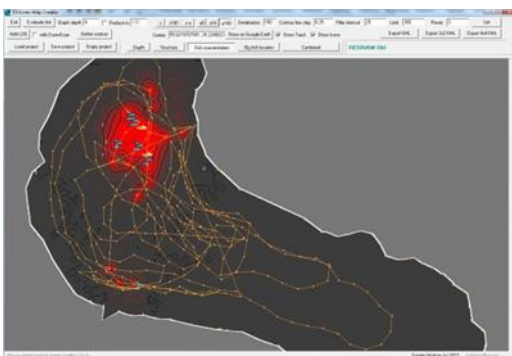
b)

Fig. 5. Map building: a) taking into account the contour of the lake; b) fish density by using CNN

V. CONCLUSION

Method for obtaining topographic maps of lakes, maps of fish concentration and a map of predator location based on the results of intelligent sonar data processing is presented. The presented algorithm is based on sonar images for the detection of classes "fish", "grass", "school of fish", "predator", "bottom fish". The algorithm includes following steps: input frame separating into overlapping blocks, blocks-processing using

CNN YOLO v2, and merging extracted bounding boxes around one object, fish concentration map building. To construct maps of the distribution of features along the lake, we propose a novel method for constructing the approximation of GPS- referenced CNN results based on the original implementation of fuzzy logic. Our method has an accuracy of 70.1% and has low percentage of false positive results in case of fish presence. To increase the accuracy we need to significantly expand the dataset for CNN training.

REFERENCES

[1] H. Balk, and T. Lindem, "Improved fish detection probability in data from split-beam sonar," in: Aquatic Living Resources, vol.13, 2000, pp.297–303. doi: 10.1016/S0990-7440(00)01079-2

[2] B. Helge, and L. Torfinn, "Improved fish detection probability in data form split-beam sonar," https://slides.tips/improved-fish-detection-probability-in-data-form-split-beam-sonar.html

[3] J. Kim, and SC. Yu, "Convolutional neural network-based real-time rov detection using forward-looking sonar image," in: Autonomous Underwater Vehicles (AUV), IEEE/OES (2016), pp. 396–400. doi: 10.1109/AUV.2016.7778702

[4] K. Krivoruchko,"Spatial Statistical Data Analysis for GIS Users," Redlands, CA: Esri Press, 2011, p.928

[5] X. Wang, "Deep Learning in Object Recognition, Detection, and Segmentation," in: Foundations and Trends in Signal Processing, vol. 8, №4, pp. 217-382

[6] R. C. Gonzalez, and R. E. Woods,"Digital Image Processing," 3rd ed., Prentice Hall, 2008, p.954

[7] C. Demant, C. Garnica, and B.Streicher-Abel, "Industrial Image Processing: Visual Quality Control in Manufacturing," 2nd ed., 2013, Springer Heidelberg, p.368

[8] Iliashenko, O. Y., Levina, A. I., Borremans, A. D. (2017). The approach to the formal specification of static structure of the system: Mapping UML to EXPRESS-G. Paper presented at the Proceedings of the 29th International Business Information Management Association Conference, 2017. P. 1177-1187.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in: NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012, vol. 1, pp. 1097-1105

[10] Sh. Ren, Kaiming He, Ross Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.39(6), 2017, pp. 1137 - 1149. doi: 10.1109/TPAMI.2016.2577031

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE Press Washington, DC, USA, 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308

[12] K. He, X. Zhang, Sh. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in: Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference, IEEE Press, Washington, DC, USA, 2016. doi: 10.1109/CVPR.2016.90

[13] Makarichev, Y. A., Anufriev, A. S., Ivannikov, Y. N., Didenko, N., & Gazizulina, A. (2018). Low - power wind generator. Paper presented at the International Conference on Information Networking, 2018-January, 671-672. doi:10.1109/ICOIN.2018.8343203

[14] J. Redmon, and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in: 2017 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Press, Washington DC 2017, pp. 6517–6525. doi:10.1109/CVPR.2017.690

[15] S. Ioffe, and Ch. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in: Proceedings of the 32nd International Conference on Machine Learning Microtome Publishing, Brookline 2015, pp. 448–456.

[16] Kozlov, A. V., Gutman, S. S., Rytova, E. V., & Zaychenko, I. M. (2017). The application of the fuzzy sets theory to valuing cumulative labor potential of the region. Paper presented at the Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017, 621-623V.

[17] Trofimov, A. Gloukhov, and D. Gloukhov, "Algorithm of ecological monitoring by fuzzy production rules," in: 2nd International Conference Ecology and Society's Development Abstracts, МАНЭБ, St.P, 1997, p.166.

[18] D. Glukhov, "Dynamic expert system by fuzzy inference rules to automations an examination of complex objects," in: Budownictwo i Inzynieria, Srodowiska, 1998, pp. 105–109.

[19] Visual Object Tagging Tool: "An electron app for building end to end Object Detection Models from Images and Videos," https://github.com/Microsoft/VoTT