

COMPUTER SCIENCE

UDC 004.932.2:004.8:681.883(045)
DOI:10.18372/1990-5548.89.21468

¹Oleksandr Katrusha,
²Kostiantyn Yefremov

A METHODOLOGY FOR MULTI-VIEW SIDE-SCAN SONAR TARGET RECOGNITION

Educational and Scientific Institute for Applied System Analysis, National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

E-mails: ¹oleksandr.katrusha@gmail.com ORCID: 0009-0008-7097-4843
²k.yefremov@wdc.org.ua ORCID:0000-0003-3495-6417

Abstract—Side-scan sonar is widely used for seabed imaging, but automated target recognition in such data remains challenging because of acoustic noise, stripe artifacts, view-dependent target appearance, navigation errors, and repeated observations of the same physical object during different passes or missions. The aim of this paper is to develop a methodology for multi-view target recognition in side-scan sonar data that enables a transition from individual sonar images and detections to the level of physical targets on the seafloor. The proposed methodology integrates mission data preparation, inter-mission change detection, estimation of the number of physical targets, detection clustering, localization, visual similarity modeling, deep model training, and multi-view classification. The input data include side-scan sonar missions, trajectory data, sonar images, metadata, and object detections, while the output consists of information about inter-mission changes, hypotheses about physical targets, their coordinates, and class labels. Individual components of the methodology were evaluated on synthetic and real datasets using task-specific quality metrics. The results showed improved robustness of change detection on noisy sonar images, estimation of the number of physical targets with an error not exceeding one object, improved clustering and localization due to the integration of spatial, topological, and visual information, and higher quality of object-level multi-view classification. The methodology forms a unified modular workflow for side-scan sonar data analysis, in which the output of one stage can be used as input or as a constraint for subsequent stages. It can serve as a basis for automated sonar data analysis systems for underwater monitoring, object search, humanitarian demining, infrastructure inspection, and operator decision support.

Keywords—Artificial intelligence, automatic target recognition, intelligent sonar data analysis, multi-view classification, computer vision, side-scan sonar.

I. INTRODUCTION

Side-scan sonar (SSS) is widely used for underwater survey missions due to its ability to acquire acoustic images of large seabed areas in turbid waters. In the current Ukrainian context, one of its key applications is underwater demining. Because of Russia’s armed aggression, Ukraine faces the need to survey about 14,000 km² of water areas, including coastal zones, rivers, and estuaries with high levels of siltation [1].

In a typical SSS mission, the same physical object may be observed multiple times, while its appearance may differ significantly depending on the vehicle trajectory and acquisition conditions. Thus, mission logs may contain hundreds of variable target detections with imprecise localization. Moreover, objects may appear or disappear between missions, further complicating the task of their recognition and localization.

At the same time most existing approaches address target recognition tasks in isolation: object

detection or classification in individual images [2], [3], [4], sonogram matching for inter-mission change detection [5], [6], detections clustering [7], [8], [9], or multi-view classification [10], [11], [12]. In addition, the development of autonomous surface and underwater vehicles leads to the generation of large amounts of mission data, which increases the workload of sonar operators and creates a need for complex automated data processing using state-of-the-art intelligent methods.

Recent research advances reveal strong potential for transferring methods and technologies from hydroacoustics, neural networks, underwater robotics, and related fields to underwater automatic target recognition (ATR) problems [1]. Therefore, there is a need to develop new methods and adapt existing approaches to the underwater acoustic domain. However, existing approaches still lack a unified end-to-end methodology that links these methods into a coherent workflow for target-level analysis of multi-mission SSS data.

The aim of this work is to develop a methodology that provides a consistent transition from a set of individual SSS images and mission metadata to hypotheses about physical seabed objects, their classes, spatial positions, and inter-mission changes. The methodology combines inter-mission change detection, estimation of the number of targets, detection clustering, localization, and multi-view classification within a single interconnected processing workflow. The methods comprising the methodology were experimentally verified on synthetic and real SSS mission datasets using task-specific quality metrics. It can serve as a basis for developing automated sonar data analysis systems for real-world applications.

II. LITERATURE REVIEW

Recent literature reviews on automated analysis of remote sensing and sonar data consider problems such as automatic target recognition, image segmentation and classification, image matching and change detection [2], [4], [13]. A substantial part of specialized research focuses on well-developed methods for ATR in individual sonar images [3], [14], [15], [16], [17]. At the same time, several studies attempt to exploit information from different sonar views to improve target classification [10], [12], [18], [19]. However, most existing works do not explicitly address the problem of grouping (clustering) individual target detections required for multi-view classification, often treating this step as a manual task. They also rarely consider changes in the set of objects between missions and their influence on subsequent ATR stages.

Classical clustering methods, including hierarchical agglomerative clustering and density-based methods such as DBSCAN, can be applied to spatial grouping of target detections in remote sensing tasks [9], [20], [21], [22]. However, their direct application to SSS data is limited, since they usually do not consider the topology of survey passes, navigation uncertainty, changes in the set of objects between missions, or visual similarity of detected objects. Besides, determining the optimal number of clusters is not only a mathematical problem of selecting a clustering parameter, but also a domain-specific problem of estimating the number of physical targets within a mission.

As for change detection in remote sensing data, pixel-based, object-based, and feature-based change detection methods based on keypoints matching are widely adopted [23], [24], [25]. However, the application of modern convolutional and transformer-based change detection methods to SSS imagery remains underexplored [6], [26], [27], [28].

Thus, the analysis of recent studies shows that individual components of automated SSS data processing for target recognition are developed unevenly and are mostly considered in isolation. This motivates the development of new methods and a methodology that integrates them into a unified SSS target recognition workflow.

III. PROBLEM STATEMENT

Consider a set of SSS missions ordered by their start time:

$$[M = M_1, M_2, \dots, M_N],$$

where each mission (M_i) contains the trajectory data of the carrier platform, a sequence of registered sonar pings, and a set of automatically or manually labelled object detections:

$$[D_i = d_{i1}, d_{i2}, \dots, d_{in_i}],$$

where (d_{ij}) denotes an individual object detection.

Each detection includes geographical coordinates of its bounding box vertices, a sonogram fragment, sonar metadata like slant range, Euler angles and additional features.

As most physical targets may be observed multiple times during different passes or missions, it is necessary to transition from a set of independent detections to a set of physical objects:

$$[O = O_1, O_2, \dots, O_K],$$

where (K) is the estimated number of physical targets observed during the missions. Each object (O_K) is associated with a subset of detections D^k belonging to the same physical target, localization coordinates (L_K), and a class label (y_k).

Then the general problem of target recognition of multi-mission SSS data can be formulated as constructing the mapping:

$$[M, D_1, D_2, \dots, D_N \rightarrow (O_k, L_k, y_k)_{k=1}^K].$$

Within this general problem, several interconnected subproblems can be identified.

The **first** subproblem is change detection between same-track missions. For pairs of missions, it is necessary to identify corresponding seabed areas, perform geometric co-registration of sonar images, and obtain information about local changes. The outputs of this stage may include raster change maps and symbolic descriptions of new or disappeared objects between missions.

The **second** subproblem is estimating the number of physical targets (K) in the observed area. Unlike the classical problem of selecting the optimal number of clusters (e.g. Silhouette method [29]), this task requires estimating the number of real seafloor objects while accounting for survey-pass topology, the spatial distribution of detections, navigation uncertainty, and the possibility of repeated observation of the same target across different missions.

The **third** subproblem is clustering the complete set of detections $[D = \bigcup_{i=1}^N D_i]$ into subsets $[C = C_1, C_2, \dots, C_K]$, where each cluster (C_K) corresponds to one physical target. During clustering, it is necessary to consider not only the spatial proximity of detections, but also constraints derived from the mission structure, information about new and disappeared objects, and the visual similarity of sonar image fragments.

The **fourth** subproblem is target localization. For each formed cluster (C_K), the coordinates of the corresponding physical target (L_K) must be estimated. For real-world usage it is also useful to estimate the spatial spread of detections to assess the reliability of the result.

The **fifth** subproblem is multi-view target classification based on the clustering results. Therefore, it is necessary to construct a classifier that takes unordered variable-size sets of images corresponding to one cluster (C_K) and determines the class of the physical target (y_k). Such multi-view classification approach can improve classification quality, as demonstrated in related works.

Thus, the solution of the stated interrelated problems requires a methodology that defines the sequence of processing stages, the relationships between the outputs of individual methods, and a system of metrics for evaluating the quality of each stage.

IV. PROPOSED METHODOLOGY

The methodology describes several stages that solve the stated problems using a mix of existing and newly developed state-of-the-art methods. Its modular architecture provides flexibility and enables continuous improvement of the methods used without redesigning the entire system. For example, the clustering method can be replaced by another detection grouping algorithm, the visual similarity model by another deep learning architecture, and so on. Besides, the methods for sonar signal and image preprocessing and generation, object detection, and

basic image keypoint matching are already well developed and can be reused in the proposed setup.

The main methodology stages and their relationships are illustrated by the schema in Fig. 1 where white boxes represent existing methods and blue boxes represent newly developed methods.

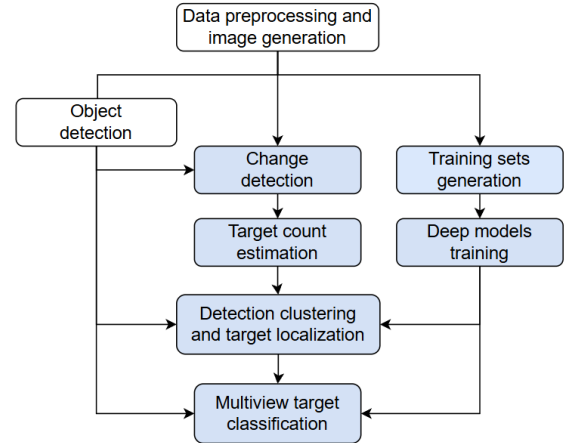


Fig. 1. Methodology schema

The general methodology comprises the following stages: data preprocessing and image generation (1), object detection (either manual or automated) (2), change detection (pixel and symbolic) between same-track missions (3), target count estimation (4), training datasets generation (5), deep models training (6), detections clustering and target localization (7), and multi-view target classification (8). The output of each stage can be used as input or as a constraint for subsequent stages, enabling effective solution of the stated problems.

At the data preprocessing and image generation stage, mission data logs are collected from vehicles, converted into a suitable format, initially reviewed by an operator, normalized by intensity, processed for water-column detection, and grouped by mission location using tools such as SonarWiz [30]. SSS ping arrays are converted into images, and mission metadata is exported to the formats suitable for further processing.

Next, object detection is performed either manually by the operator or automatically by modern deep models and detection coordinates are transformed into a common local metric coordinate system. The image patches of the detected regions of interest and related metadata (like estimated geocoordinates of the bounding box points) are passed to the next stages.

The change detection stage analyzes differences between same-track missions. The missions are split into relatively straight segments (tracklines) by geographical proximity. For the corresponding sonar

images, keypoints matching and geometric alignment are performed in a common coordinate frame using convolutional and transformer-based matching models [26]. Based on the registered images, raster change maps can be generated to represent local intensity differences, while symbolic change maps describe new and disappeared detections. This stage serves both as an independent change analysis tool and as a source of additional constraints for subsequent detection clustering.

Target count estimation accounts for survey-pass topology, spatial distribution of detections, probability of repeated target observation, navigation uncertainty, and possible changes in the set of objects between missions. The estimation is performed using Monte Carlo simulation of detection distribution and Bayesian inference of the most probable number of targets. The resulting target count is then used as a parameter for clustering.

The central stage of the methodology is detection clustering. It is based on a modified agglomerative clustering algorithm and considers the estimated number of targets, spatial distances between detections, survey-pass topology and object distribution, information about new and disappeared objects, and, when available, visual similarity features.

To enable effective use of modern deep models in subsequent stages, a custom training dataset is constructed. It contains a target identifier for each image patch and thus enables target-oriented model training.

To improve clustering, a deep visual similarity triplet model with a target-oriented algorithm of triplet generation is introduced [31]. It forms vector representations of patches and allows the cosine distance between patches to be calculated and used as an additional component of the clustering distance metric. An example of detection clustering is shown in Fig. 2 as a screenshot from the implemented software system. Gray triangles denote object detections, while colored squares indicate estimated targets.

After clustering, target localization is performed. For each cluster, the position of the corresponding physical target is estimated as the mean of its detection coordinates. Additional spatial spread statistics, such as the number of detections per cluster and the standard deviation of detections within a cluster, help assess localization reliability, identify outliers, and justify spatial thresholds and clustering hyperparameters.

The final stage is multi-view target classification based on the clustering results. Since the number of observations may vary and their order has no physical meaning, permutation-invariant aggregation

methods are appropriate. In this methodology, a multi-view classifier based on the Deep Sets architecture [32], which reuses the convolutional object detector backbone, is proposed to improve robustness to noise, view dependence, and ambiguity of single sonar images [10].

Together, localization and multi-view classification complete the transition from clustered detections to object-level interpretation of the underwater scene. Thus, the proposed methodology forms a unified SSS data analysis workflow for multi-view target recognition. Table I illustrates the stages, their input and output data, and their role in the methodology. It also shows how intermediate outputs are transferred between modules and used to support subsequent processing stages.

The proposed methodology was implemented as part of a software system for sonar mission analysis. The software integrates the main processing modules of the methodology, including mission data preparation, inter-mission image matching and change detection, target count estimation, detection clustering, localization, visual similarity estimation, and multi-view classification. Therefore, the software system serves not only as an experimental platform for validating the developed methods, but also as a prototype of an automated system for multi-view target recognition in SSS data.

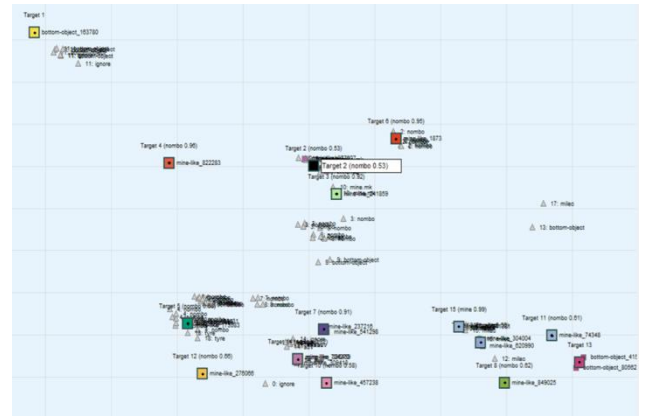


Fig. 2. An example of detection clustering results of a real SSS mission

V. EVALUATION FRAMEWORK

The methodology evaluation requires task-specific metrics for each processing stage. For image matching and change detection, the main indicators include the number of keypoint matches, the percentage of geometrically consistent correspondences after RANSAC filtering, mean reprojection error, spatial displacement, matching angle stability, computational cost, runtime, and memory consumption [26].

TABLE I. STAGES OF THE PROPOSED METHODOLOGY

Stage	Input data	Output data	Methods and tools
Data preprocessing and image generation	SSS mission logs	SSS images and converted mission metadata	Industry-adopted software like SonarWiz and custom algorithms
Object detection	Waterfall SSS images	List of detections with pixel coordinates	Modern convolutional object detectors like YOLO family; Industry-adopted software like SonarWiz
Change detection	SSS images and mission metadata	Pixel and symbolic change maps	Convolutional and transformer keypoints detectors; custom change detection method
Target count estimation	Mission metadata, symbolic change maps	Estimated number of targets	Monte-Carlo simulation and Bayesian inference
Training sets generation	SSS waterfall images and mission metadata, ground truth about target classes and location	Dataset with target identifier for each patch	Custom method for dataset generation
Deep models training	Dataset with target identifier for each patch	Visual similarity model; Deep Sets-based classification model	Custom triplet generation method
Detection clustering and target localization	Detections, estimated number of targets, constraints, cosine similarity distance	Detection clusters, Estimated target locations; clustering report	Modified agglomerative clustering; triplet visual similarity model
Multi-view target classification	Sets of image patches for each cluster	Estimated target class	Deep Sets or other permutation-invariant aggregation models

Target count estimation is evaluated by comparing the estimated number of physical targets with the ground truth number of objects in the observed area. The metrics may include absolute error and the proportion of cases in which the error does not exceed a predefined threshold, for example one object.

Detection clustering is evaluated by comparing the obtained clusters with ground truth detection groups corresponding to physical targets. The main evaluation metrics include F1-score and adjusted Rand index, which characterize both pairwise grouping quality and overall agreement between the obtained and reference clustering.

Target localization is evaluated by comparing estimated target coordinates with ground truth positions using root mean square error (RMSE).

For visual similarity estimation, metrics should assess the quality of the learned latent space, including distances between detections of the same and different physical targets and the impact of visual similarity on clustering performance.

For multi-view classification, standard classification metrics are used, including accuracy, precision, recall, and F1-score. Results can be assessed both at the patch level and at the target level, with the latter being primary in the proposed methodology.

Thus, the evaluation framework links each methodology stage with relevant quality indicators,

enabling analysis of the final recognition result, the contribution of intermediate stages, error sources, and the overall effectiveness of the proposed multi-view SSS target recognition workflow.

VI. EXPERIMENTAL RESULTS

Each method proposed in the methodology was tested using task-specific datasets, selected baseline methods, and evaluation metrics.

The change detection method was evaluated on pairs of sonar images acquired under different conditions. On the noisy Atlantic dataset, deep matching models increased the RANSAC inlier percentage by 16.2 percentage points and reduced the mean reprojection error by a factor of 15.7 compared with baseline methods. On the clearer Baltic dataset, the improvement was smaller: the RANSAC inlier percentage increased by 4.3 percentage points, while the reprojection error decreased by about 1.2 times. Dense transformer-based methods also increased the number of matches up to 19 times, which is important for image registration and change map generation.

The target count estimation method was tested on datasets with known numbers of physical objects. The method estimated the number of physical targets with an error not exceeding one object, which is important for constraining subsequent clustering and reducing both over-splitting and false merging.

Detection clustering and localization were evaluated by comparing generated clusters with reference groups of physical targets. The baseline included agglomerative, k-means and DBSCAN clustering based on Euclidean distance between detection center coordinates. Incorporating domain-specific constraints improved grouping quality: the clustering F1-score increased by 9–21 percentage points, while the adjusted Rand index increased by 6–18 percentage points. Improved clustering also increased localization accuracy, with RMSE reduction reaching 9%.

The use of a weighted metric combining spatial distance and visual similarity improved clustering compared with spatial features alone, increasing the F1-score by up to 10 percentage points. This confirms that visual information helps distinguish nearby objects and associate observations of the same target under navigation uncertainty.

The multi-view classification method was evaluated on sets of detections grouped by physical targets. Object-level multi-view classification outperformed single-view and ensemble approaches, with macro F1-score increasing by up to 24 percentage points compared to baseline.

Overall, the experimental results confirm the effectiveness of the proposed methodology components and support their integration into a unified multi-view SSS target recognition workflow. The obtained improvements in change detection, target count estimation, clustering, localization, visual similarity integration, and object-level classification demonstrate that considering spatial, topological, visual, and multi-view information jointly provides more reliable interpretation of multi-mission sonar data than isolated processing of individual images or detections.

The main limitations are related to the quality of real SSS data. Errors in initial detections, missed objects, false alarms, and inaccurate navigation can affect target count estimation, clustering, localization, and multi-view classification. Another limitation is the difficulty of constructing fully annotated end-to-end datasets containing repeated missions, object-level labels, detection-to-target correspondences, change information, and class labels.

VII. CONCLUSIONS

The paper proposes a methodology for multi-view side-scan sonar target recognition. It integrates inter-mission change detection, estimation of the number of physical targets, detection clustering, localization, deep model training, and multi-view classification into a unified processing workflow that enables a

transition to the level of physical seabed objects and accounts for SSS data peculiarities.

A generalized problem statement is formulated in which the input data include SSS missions' metadata, sonar images, and detections, while the output consists of change information, physical targets, their estimated coordinates, and class labels. The methodology defines the sequence of processing stages, their inputs and outputs, and the methods and tools proposed for use at each stage.

An evaluation framework is proposed for the main stages of the methodology. Experimental results for individual methodology components demonstrate improvements in the quality of the corresponding tasks and support the feasibility of the proposed approach.

The practical significance of the methodology lies in its applicability as a basis for automated SSS data analysis systems for underwater monitoring, object search, humanitarian demining, infrastructure inspection, and operator decision support. Its modular structure allows individual components to be used independently or as part of a unified workflow.

Future work should focus on creating fully annotated multi-mission SSS datasets, studying error propagation between stages, integrating navigation refinement and SLAM, and developing multi-view classifiers for a wider range of underwater object classes.

REFERENCES

- [1] O. Drobot, "Underwater development as a scientific direction: integration of innovation, analytics and academic potential," presented at the National Conference on Underwater Development "Safety Depth," Ministry defense of Ukraine, Nov. 2025. [in Ukrainian].
- [2] M. Motylinski, A. J. Plater, and J. E. Higham, "Computer vision methods for side scan sonar imagery," *Meas. Sci. Technol.*, vol. 36, no. 1, p. 015435, Jan. 2025, <https://doi.org/10.1088/1361-6501/ad99f1>.
- [3] Z. Heng, H. Shuping, X. Jingfeng, H. Yaohui, and H. Yubo, "A Review of Intelligent Detection Methods for Underwater Targets in Sonar Images," in *2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Sep. 2023, pp. 1462–1468. <https://doi.org/10.1109/ITOEC57671.2023.10291340>.
- [4] Y. Steiniger, D. Kraus, and T. Meisen, "Survey on deep learning based computer vision for sonar imagery," *Engineering Applications of Artificial*

- Intelligence*, vol. 114, p. 105157, Sep. 2022, <https://doi.org/10.1016/j.engappai.2022.105157>.
- [5] G. Cheng *et al.*, “Change Detection Methods for Remote Sensing in the Last Decade: A Comprehensive Review,” May 09, 2023, arXiv:2305.05813. <https://doi.org/10.48550/arXiv.2305.05813>.
- [6] Y. Steiniger and S. Schröder, “Studying Deep Learning Based Co-Registration for Incoherent Change Detection,” in *International Conference on Underwater Acoustics 2024*, Bath, UK: Institute of Acoustics, May 2024. <https://doi.org/10.25144/22276>.
- [7] X. Fan, L. Lu, P. Shi, and X. Zhang, “A novel sonar target detection and classification algorithm,” *Multimed Tools Appl*, vol. 81, no. 7, pp. 10091–10106, Mar. 2022, <https://doi.org/10.1007/s11042-022-12054-4>.
- [8] S. Zhuang and C. Chen, “Aerial battlefield targets grouping based on DTW-DBSCAN algorithm,” in *2021 40th Chinese Control Conference (CCC)*, Shanghai, China: IEEE, Jul. 2021, pp. 3397–3402. <https://doi.org/10.23919/CCC52363.2021.9550653>.
- [9] L. Zhang, Q. Wang, Y. Guo, S. Xu, and L. Wang, “A Novel Method for Small-Target Detection in Sea Clutter: Spectral Clustering Based on Neighborhood Density Similarity Measure,” *IEEE Sensors J.*, vol. 25, no. 2, pp. 2988–2997, Jan. 2025, <https://doi.org/10.1109/JSEN.2024.3499954>.
- [10] O. Katrusha, “Classification of Targets from Multi-View Side-Scan Sonar Imagery,” presented at the ASPAI’26, Jun. 26, 2026. [Online]. Available: https://sensorsportal.com/ASPAI/ASPAI_2026_Proceedings.pdf
- [11] V. Myers and D. P. Williams, “Adaptive Multiview Target Classification in Synthetic Aperture Sonar Images Using a Partially Observable Markov Decision Process,” *IEEE J. Oceanic Eng.*, vol. 37, no. 1, pp. 45–55, Jan. 2012, <https://doi.org/10.1109/JOE.2011.2175510>.
- [12] D. P. Williams, R. Hamon, and I. D. Gerg, “On the Benefit of Multiple Representations with Convolutional Neural Networks for Improved Target Classification Using Sonar Data,” in *Underwater Acoustic Conference and Exhibition Series*, 2019, pp. 187–194.
- [13] A. Khan, M. M. Fouda, D.-T. Do, A. Almaleh, A. M. Alqahtani, and A. U. Rahman, “Underwater Target Detection Using Deep Learning: Methodologies, Challenges, Applications, and Future Evolution,” *IEEE Access*, vol. 12, pp. 12618–12635, 2024, <https://doi.org/10.1109/ACCESS.2024.3353688>.
- [14] X. Luo, L. Chen, H. Zhou, and H. Cao, “A Survey of Underwater Acoustic Target Recognition Methods Based on Machine Learning,” *JMSE*, vol. 11, no. 2, p. 384, Feb. 2023, <https://doi.org/10.3390/jmse11020384>.
- [15] X. Du, Y. Sun, Y. Song, H. Sun, and L. Yang, “A Comparative Study of Different CNN Models and Transfer Learning Effect for Underwater Object Classification in Side-Scan Sonar Images,” *Remote Sensing*, vol. 15, no. 3, p. 593, Jan. 2023, <https://doi.org/10.3390/rs15030593>.
- [16] Y. Chai, H. Yu, L. Xu, D. Li, and Y. Chen, “Deep Learning Algorithms for Sonar Imagery Analysis and Its Application in Aquaculture: A Review,” *IEEE Sensors J.*, vol. 23, no. 23, pp. 28549–28563, Dec. 2023, <https://doi.org/10.1109/JSEN.2023.3324438>.
- [17] L. C. F. Domingos, P. E. Santos, P. S. M. Skelton, R. S. A. Brinkworth, and K. Sammut, “A Survey of Underwater Acoustic Data Classification Methods Using Deep Learning for Shoreline Surveillance,” *Sensors*, vol. 22, no. 6, p. 2181, Jan. 2022, <https://doi.org/10.3390/s22062181>.
- [18] X. Tong and Y. Wang, “Active Multiview Fusion Framework for SAR Automatic Target Recognition,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, 2024, <https://doi.org/10.1109/TIM.2023.3341140>.
- [19] R. A. Dakhil and A. R. H. Khayeat, “Review on Deep Learning Techniques for Underwater Object Detection,” in *Data Science and Machine Learning*, Academy and Industry Research Collaboration Center (AIRCC), Sep. 2022, pp. 49–63. <https://doi.org/10.5121/csit.2022.121505>.
- [20] Q. Teng, X. Yang, Q. Sun, P. Wang, X. Wang, and T. Xu, “Sequential attention layer-wise fusion network for multi-view classification,” *Int. J. Mach. Learn. & Cyber.*, Jul. 2024, <https://doi.org/10.1007/s13042-024-02260-x>.
- [21] L. Yan *et al.*, “EM-Based Algorithm for Unsupervised Clustering of Measurements from a Radar Sensor Network,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 1, pp. 787–801, Feb. 2025, <https://doi.org/10.1109/TAES.2024.3448390>.
- [22] Y. Luo, X. Li, J. Wang, and S. Chen, “Clustering and Tracking-Guided Infrared Spatial-Temporal Small Target Detection,” *IEEE Trans. Geosci. Remote Sensing*, vol. 62, pp. 1–20, 2024, <https://doi.org/10.1109/TGRS.2024.3384440>.
- [23] Y. Zhang, P. Peng, and R. Lu, “Research on Target Grouping Method Based on Graph Neural Network,” in *2022 7th International Conference on Computer and Communication Systems (ICCCS)*, Wuhan, China: IEEE, Apr. 2022, pp. 32–36. <https://doi.org/10.1109/ICCCS55155.2022.9846007>.

- [24] A. Shafique, G. Cao, Z. Khan, M. Asad, and M. Aslam, "Deep Learning-Based Change Detection in Remote Sensing Images: A Review," *Remote Sensing*, vol. 14, no. 4, p. 871, Feb. 2022, <https://doi.org/10.3390/rs14040871>.
- [25] Y. Afaq and A. Manocha, "Analysis on change detection techniques for remote sensing applications: A review," *Ecological Informatics*, vol. 63, p. 101310, Jul. 2021, <https://doi.org/10.1016/j.ecoinf.2021.101310>.
- [26] O. Katrusha, D. Prylipko, and K. Yefremov, "Change detection in side-scan sonar imagery based on deep learning feature matching methods," *EEJET*, vol. 6, no. 2 (138), pp. 52–62, Dec. 2025, <https://doi.org/10.15587/1729-4061.2025.346940>.
- [27] W. Jiang, Y. Sun, L. Lei, G. Kuang, and K. Ji, "Change detection of multisource remote sensing images: a review," *International Journal of Digital Earth*, vol. 17, no. 1, p. 2398051, Dec. 2024, <https://doi.org/10.1080/17538947.2024.2398051>.
- [28] W. Hedlund, "Change Detection in Synthetic Aperture Sonar Imagery Using the Segment Anything Model," 2024.
- [29] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, Nov. 1987, [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [30] "SonarWiz Sidescan." Accessed: Mar. 03, 2026. [Online]. Available: <https://chesapeakeotech.com/products/sonarwiz-sidescan/>
- [31] O. Katrusha, "Deep Similarity Model for Improving Clustering of Sonar Object Detections," in *Science, Technology and Industry in the Digital Age*, Hamburg, Germany: International Scientific Unity, Mar. 2026, pp. 243–246. <https://doi.org/10.70286/isu-18.03.2026.038>.
- [32] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, and A. Smola, "Deep Sets," *Machine Learning*, Apr. 14, 2018, <https://doi.org/10.48550/arXiv.1703.06114>.

Received: July 19, 2026

Accepted: August 12, 2026

Published: September 18, 2026

Katrusha Oleksandr. ORCID 0009-0008-7097-4843. Master of Science. Postgraduate Student.

Department of Artificial Intelligence, Educational and Scientific Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

Research area: artificial intelligence, computer vision, sonar data processing

Number of publications: 5

E-mail: oleksandr.katrusha@gmail.com

Yefremov Kostiantyn. ORCID 0000-0003-3495-6417. Candidate of Technical Sciences. Associate Professor.

Director of Educational and Scientific Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

Research area: artificial intelligence, intelligent information systems, data analysis and integration, geoinformation technologies, decision support systems, sustainable development

Number of publications: 26.

E-mail: k.yefremov@wdc.org.ua

О. М. Катруша, К. В. Єфремов. Методологія багатовидового розпізнавання цілей у даних сонара бокового огляду

Сонар бокового огляду широко використовується для отримання зображень дна водойм, однак автоматизоване розпізнавання цілей у таких даних ускладнюється акустичним шумом, смуговими артефактами, залежністю вигляду об'єкта від ракурсу спостереження, похибками навігації та повторним виявленням однієї фізичної цілі у різних проходах або місіях. Метою роботи є розроблення методології багатовидового розпізнавання цілей у даних сонара бокового огляду, яка забезпечує перехід від окремих сонарних зображень і детекцій до рівня фізичних об'єктів на дні. Запропонована методологія поєднує підготовку даних місій, виявлення змін між повторними місіями, оцінювання кількості фізичних цілей, кластеризацію детекцій, локалізацію, формування ознак візуальної подібності, навчання глибоких моделей та багатовидову класифікацію. Вхідними даними є місії сонара, траєкторні дані, сонарні зображення, метадані та детекції об'єктів, а вихідними результатами – гіпотези про фізичні цілі, їх координати, інформація про зміни між місіями та класові мітки. Окремі складові методології перевірено на синтетичних і реальних наборах даних з використанням відповідних метрик якості. Результати показали підвищення стійкості виявлення змін на зашумлених зображеннях, оцінювання кількості цілей з похибкою не більше одного об'єкта, покращення кластеризації та локалізації за рахунок просторових, топологічних і візуальних ознак, а також підвищення якості багатовидової класифікації на рівні фізичних об'єктів. Методологія формує єдиний модульний процес аналізу багаторазових місій сонара бокового огляду та може бути використана як основа для автоматизованих систем підводного моніторингу, пошуку об'єктів, гуманітарного розмінування, обстеження інфраструктури та підтримки прийняття рішень оператором.

Ключові слова: автоматичне розпізнавання цілей, багатовидова класифікація, інтелектуальний аналіз сонарних даних, комп'ютерний зір, сонар бокового огляду, штучний інтелект.

Катруша Олександр Миколайович. ORCID 0009-0008-7097-4843. Магістр. Аспірант.

Кафедра Штучного інтелекту, Навчально-науковий інститут прикладного системного аналізу, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Напрямок наукової діяльності: штучний інтелект, комп'ютерний зір, обробка сонарних даних.

Кількість публікацій: 5.

E-mail: oleksandr.katrusha@gmail.com

Єфремов Костянтин Вікторович. ORCID 0000-0003-3495-6417. Кандидат технічних наук. Доцент.

Директор Навчально-наукового інституту прикладного системного аналізу. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Напрямок наукової діяльності: штучний інтелект, інтелектуальні інформаційні системи, аналіз та інтеграція даних, геоінформаційні технології та системи підтримки прийняття рішень, сталий розвиток.

Кількість публікацій: 26.

E-mail: k.yefremov@wdc.org.ua