Multi-Modal Obstacle Avoidance in USVs via Anomaly Detection and Cascaded Datasets[*]

Tilen Cvenkel, Marija Ivanovska, Jon Muhovič, and Janez Perš[0000−0002−6039−6110]♠

Faculty of Electrical Engineering
University of Ljubljana, Tržaska 25, Ljubljana, Slovenia
♠ Corresponding author: `janez.pers@fe.uni-lj.si`
https://lmi.fe.uni-lj.si/en/home/

Abstract. We introduce a novel strategy for obstacle avoidance in aquatic settings, using anomaly detection for quick deployment of autonomous water vehicles in limited geographic areas. The unmanned surface vehicle (USV) is initially manually navigated to collect training data. The learning phase involves three steps: learning imaging modality specifics, learning the obstacle-free environment using collected data, and setting obstacle detector sensitivity with images containing water obstacles. This approach, which we call *cascaded datasets*, works with different image modalities and environments without extensive marine-specific data. Results are demonstrated with LWIR and RGB images from river missions.

Keywords: unmanned vehicles · USV · obstacle avoidance · anomaly detection

1 Introduction

Obstacle detection is crucial for autonomous vehicles. Unmanned robotic surface vehicles (USVs) can use various sensors, like RGB cameras [37], sometimes in stereo depth configuration [32], RADAR [35], LIDAR [23, 41], and SONAR [20]. Cameras are appealing for their cost and superficial similarity to human perception, but require substantial image processing, which falls in the domain of *computer vision*.

Cameras' main drawback is their sensitivity to environmental variations. However, lately data-driven algorithms and deep neural networks (DNN) improved obstacle detection in marine environments [5, 29, 33], but require extensive annotations [46].

We propose *semi-supervised learning* for water obstacle detection, specifically *anomaly detection in a one-class learning setting* [42, 3, 39]. This approach trains on normal data and detects anomalies as non-conforming samples. Many, (but not all) USVs are expected to operate in limited geographical domains, allowing one-class learning and anomaly detection.
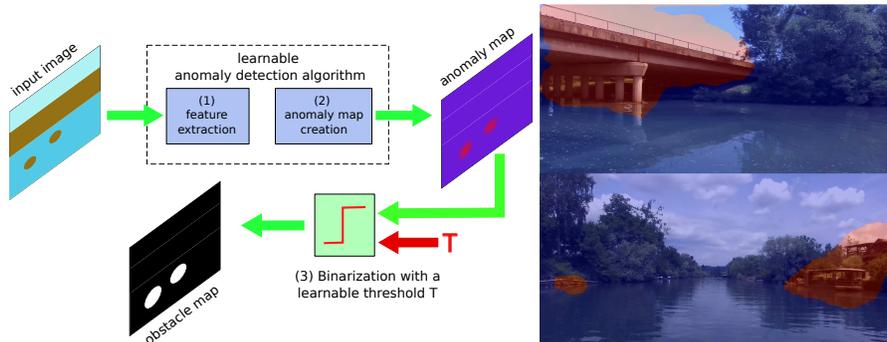
Fig. 1. Left: Obstacle avoidance through anomaly detection has three learnable components [(1), (2), (3)]. This approach allows separate learning using datasets without marine environments (1), obstacles (2), and with minimal detailed obstacle annotations (3). Right: Obstacle detection examples using one-class learning anomaly detector CS-Flow [39]: in the right-upper image correctly detects a bridge, while in the right-lower image detects a floating dock and a riverbank house.

Our contributions include: *i)* a novel strategy for water obstacle detection with limited annotated data, *ii)* a new approach to training with limited marine data, *iii)* evaluation of recent SOTA anomaly detection algorithms in realistic USV scenarios, and *iv)* comparison of a fully-supervised SOTA segmentation algorithm [6] with our approach.

2 Related work

We review works using image data and computer vision for water obstacle detection, followed by literature on semi-supervised learning for anomaly detection, which avoids laborious annotation.

2.1 Detection of water obstacles

Detecting water obstacles under varying weather conditions is challenging. Early research used hand-crafted methods [49, 22], while SOTA approaches leverage deep learning for more discriminative models. Many works [28, 31, 51] propose CNNs for vessel detection and classification. However, these don't generalize well and fail to recognize arbitrary objects.

Recent algorithms use neural networks for semantic segmentation. Cane *et al.* [10] and Bovcon *et al.* [7] applied segmentation models to marine data, while Kim *et al.* [26], Zhan *et al.* [55], and Steccanella *et al.* [43] proposed adapted segmentation architectures. The SOTA method [6] estimates water horizon location, fusing inertial information with RGB data. A time-efficient model is presented in [52]. Segmentation-based models require massive annotated data, posing a limitation, especially for less popular image modalities [33].

2.2 Anomaly detection in one–class learning setting

Anomaly detection in industrial applications motivated one-class learning algorithms, trained using only normal data samples [40, 42, 1]. Reconstruction-based deep learning models use auto-encoder architectures, improved by pretext learning objectives [15, 34, 54, 19, 2, 21, 45].

Recent one-class anomaly detectors utilize pre-trained classification networks like VGG16, ResNet, and EfficientNet as backbone models. Defard *et al.* [13] proposed modeling pre-extracted features using a multivariate Gaussian distribution. Others pursued shallow approaches [38, 11], knowledge distillation [30, 4, 47], or invertible neural networks [39, 18, 53] to construct flexible probability distributions of non-anomalous data.

2.3 Multi-modal data

Autonomous navigation needs multi-modal data to perform well in diverse conditions, but the research into this problem is not as advanced as in RGB modality. Nirgudkar and Robinette [33] use LWIR modality. Datasets like nuScenes [9], Cityscapes [12], KITTI [16], and Waymo Open Dataset [50] include multiple RGB, depth cameras, lidar, and GPS. Pedestrian detection datasets like LLVIP [24] and CVC-14 [17] use LWIR and RGB cameras for better low-light performance. Wang et al. [48] propose a method for anomaly detection in hyperspectral satellite images.

3 Extension of anomaly detection to water environment

The presented solution to the obstacle detection problem as shown in Figure 1 is based on a cascade of training datasets.

3.1 Cascaded datasets

Our assumption is that the following non-overlapping image datasets are at our disposal:

- **Modality adaptation dataset**: Used for training a deep CNN architecture using a proxy task (e.g. general object detection on ImageNet). No pixel-wise annotations needed. Used in stage (1) of our approach, as in Figure 1.
- **Environment adaptation dataset**: Trains the obstacle detector to the obstacle-free environment. No data annotations needed. Images must match the modality of the image acquisition hardware. No pixel-wise annotations needed. Used in stage (2) of our approach, shown in Figure 1.
- **Tuning dataset**: Fine-tunes the model's sensitivity to obstacles using pixel-wise annotated images. We used 32 image items. The detection threshold T balances true and false positives. Tuning dataset is part of the stage (3) in Fig. 1.

3.2 Tuning process and tuning targets

Anomaly detection algorithms yield prediction maps $f(u, v)$, with higher values indicating a higher certainty of anomalies. Binarization threshold T determines the operating point of the whole system and makes the it more or less sensitive to anomalies.

Evaluation and T tuning metrics depend on the detection model's functionality. While F1-score is widely used for comparing algorithms [8, 25], FNR (the false negative rate – the probability that the USV runs over an obstacle) and FPR ction algorithms, FNR (the false negative rate – the probability that the USV runs over an obstacle) and FPR (the false positive rate – the probability that USV is unnecessarily stopped) are more intuitive. Detection goals can change based on the USV's location. In our approach, varying threshold T according to a function depending on the vehicle's GPS position could achieve different mission goals.

4 Experiments

We based our research on the LWIR and RGB image data, that was acquired using our own USV multi-sensor system [36], attached to a river boat. RGB and LWIR images were recorded. [1] The image acquisition took place on a stretch of Ljubljanica river [2], that represents predominantly natural (river,bush) environment with some urban elements.

4.1 Dataset

Data was gathered in June 2021 and September 2021. In June, only RGB images were taken under good weather conditions, while in September, both RGB and LWIR images were collected under cloudy conditions.

Data was organized into three main datasets: *LWIRSept*, *RGBJune*, and *RGBSept*, which were further divided into smaller subsets. These three datasets were then further divided into 5 smaller, non-overlapping subsets, i.e. *LWIRSeptemberNoObs*, *RGBJuneNoObs*, both consisting of obstacle–free images and *LWIRSeptObs*, *RGBJuneObs*, *RGBSeptObs*, containing images with obstacles. Tuning datasets, *LWIRSeptObs32* and *RGBJuneObs32*, were created using selected images (see subsection 3.1). An overview of the datasets is presented in Table 1. Images in some subsets don't depict the same geographical location; subsets with obstacles and without obstacles were obtained on different river stretches.

In our experiments, we use *LWIRSeptNoObs* and *RGBJuneNoObs* in the role of an environment adaptation dataset (see subsection 3.1), and *LWIRSeptObs32*, *RGBJuneObs32* and *RGBSeptObs33* as tuning datasets. and *LWIRSep-*

---

[1] Stereolabs ZED stereo camera (only the left frame) and Device A-lab SmartIR384L thermal camera

[2] Data was sampled from a section between 46.0402°N, 14.5125°E and 46.0234°N, 14.5079°E

Table 1. Structure of our experimental dataset

Subset	Images	Annotated	Obstacles	Purpose
LWIRSeptNoObs	3263	No	No	Environment adaptation
LWIRSeptObs	436	Yes	Yes	Testing
LWIRSeptObs32	32	Yes	Yes	Parameter tuning
RGBJuneNoObs	600	No	No	Environment adaptation
RGBJuneObs	501	Yes	Yes	Testing
RGBJuneObs32	32	Yes	Yes	Parameter tuning
RGBSeptObs	521	Yes	Yes	Testing
RGBSeptObs33	33	Yes	Yes	Parameter tuning

tObs, *RGBJuneObs* and *RGBSeptObs* for the final evaluation of the trained obstacle detection model. In the testing phase of the trained model, *RGBJuneObs* was used to evaluate the algorithm under same weather conditions. *RGBSeptObs* was on the other hand obtained under very different weather conditions, so this dataset is used to evaluate the detection accuracy of the model in a much more challenging scenario. Note, that our acquired data (described in Table 1) does not provide any modality adaptation datasets. Since such images are not necessarily domain–specific, any publicly available dataset, representing the imaging modality of interest, can be used in this role. In our experiments ImageNet [14] was used as a modality adaptation dataset for the obstacle detection in RGB images, while Teledyne FLIR Thermal Dataset [44] was used for obstacle detection in LWIR images.

4.2 Evaluation

Our evaluation protocol is based on several assumptions, that are realistic for USV environments and commonly used in such scenarios. We consider the performance of the algorithm in the upper part of the image entirely irrelevant, as this part of the image contains the sky, and possibly the distant shore [27]. The demarcation line between the upper and bottom part of the image can be inferred from the inertial sensor in the vehicle, which has been shown to help with image segmentation in marine environment before [6]. Due to the slow dynamic of the vessel used for image acquisition (imperceptible pitch and roll), we evaluate our algorithms using a fixed horizontal line, as shown in Figure 2.
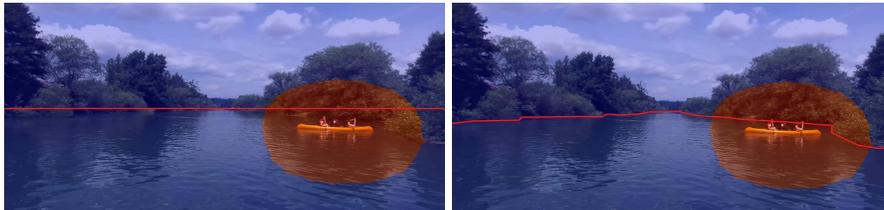

Fig. 2. Left: Simplified water edge annotation, straight line. Right: precise polygon-based annotation of the water edge. The former could be easily inferred from USV's inertial sensor, as proposed by [6]. Our method is evaluated using both kinds of annotations.

In the testing phase, only the part of the image below the water edge is used for the quantitative evaluation of the trained model. Interesting structures, such as bridges and riverbank houses, appearing above the water edge are evaluated only qualitatively. Two such examples are presented in the Figure 1. To exclude all detections on the riverbank, we perform additional evaluation, using more detailed polygon annotations.

We evaluated two different recent anomaly detection algorithms, i.e. PaDiM [13] and CS-Flow [39]. PaDiM was tested on both, LWIR and RGB images, whereas CS-Flow was evaluated only on RGB images.

4.3 Training

The modality adaptation stage for the RGB images was skipped in the publicly available models of PaDiM and CS-Flow, since they both use a backbone classification CNN, pre–trained on RGB images. The modality adaptation was thus performed only for LWIR images. We trained feature extraction CNN from scratch, using the Teledyne FLIR Thermal Dataset for Algorithm Training [44] in object detection task as a proxy.

For the environment adaptation stage, PaDiM was trained on *RGBJuneNoObs* and *LWIRSeptNoObs* subsets, to adapt it to the target river environment for both (RGB and LWIR) camera modalities, respectively, resulting in two different models, one for each imaging modality. CS-Flow was trained on *RGBJuneNoObs* only, to adapt it to target river environment in the RGB images.

Finally, in the tuning stage, thresholds T were obtained using MODS evaluation scheme [8] on *RGBJuneObs32* for RGB models, and *LWIRSeptObs32* for the PaDiM model, which is adapted to LWIR images. Final threshold values were selected based on the highest $F1$ score.

Comparison to state–of–the–art To provide a comparison of our method with the fully–supervised SOTA segmentation algorithms, we used the publicly available version of WaSR [6], pre–trained on RGB images. The algorithm was used as an out–of–the–box method, without any modifications. Since [33] shows that WaSR does not work on LWIR images without retraining, we did not use it on LWIR.

5 Results

Each evaluation was performed twice, using first the straight line as annotation of water boundary, and then for the second time, using a more accurate polygon to delimit the water area where the evaluation is performed, as shown in Figure 2.

5.2 Testing in significantly different weather conditions

5.1 Testing in similar weather conditions

Both RGB adapted algorithms were then also evaluated on *RGBSeptObs*, where the weather conditions differ significantly from the *RGBJuneNoObs* training dataset.

Both RGB models, PaDiM and CS-Flow, were tested on *RGBJuneObs* subset, while LWIR PaDiM was tested on *LWIRSeptObs*. Obtained results are reported in Tables 2 and 3.

Table 2. PaDiM results on water edge annotated as polygon.

Environment adaptation	Tuning	Testing	TP	FP	FN	F1
LWIRSeptNoObs	LWIRSeptObs32	LWIRSeptObs	496	286	399	59.2%
RGBJunNoObs	RGBJunObs32	RGBJunObs	614	270	272	69.4%
RGBJunNoObs	RGBJunObs32	RGBSeptObs*	38	40	1529	4.6%

* denotes different weather conditions

Table 3. PaDiM results on water edge annotated with fixed horizontal line.

Environment adaptation	Tuning	Testing	TP	FP	FN	F1
LWIRSeptNoObs	LWIRSeptObs32	LWIRSeptObs	692	267	279	71.7%
RGBJunNoObs	RGBJunObs32	RGBJunObs	1116	613	565	65.5%
RGBJunNoObs	RGBJunObs32	RGBSeptObs*	52	40	2260	4.3%

* denotes different weather conditions

worse than the results obtained under similar weather conditions. Obtained metrics values are reported in Tables 2, 3, 4 and 5.

For each of the previously described testing scenarios we report the number of true positives (TP), true negatives (TN), false positives (FP), and F1-score. Finally, in Table 6 we provide results of the SOTA segmentation–based algorithm WaSR, [6], for comparison.

From the results it can be concluded that WaSR still outperforms both anomaly detection methods in terms of the chosen evaluation metrics. If we compare these two methods to one another, we can see, that the CS-Flow [39] outperforms PaDiM [13], where the distribution of data features is modelled using a strong statistical prior. The difference is especially striking in situation where the training data is captured in sunny weather, while the testing data is captured in the rainy weather (Tables 2,3,4,5).

However, this comparison is not entirely fair due to different data that has gone into training WaSR on RGB images. It needed 1320 accurately pixel-wise annotated training images [7], and efforts to improve its generalization with further training with diverse images from all around the world are still ongoing. This will be difficult to repeat with many other, less widespread modalities, such as near-infrared (NIR), LWIR (done only on geographically limited area so far), and especially various multi-spectral cameras.

To complete the insight into the performance of the anomaly detection methods on our data, we present visual evaluation on two images, containing above-water obstacles, as shown in Figure 1. As can be seen, these particular examples show, that the methods successfully recognize any objects, that were not part of domain adaptation subset, as an anomaly, which is the key advantage of the one–class learning approach.

**Table 4.** CS-Flow results on water edge annotated as polygon.

Environment adaptation	Tuning	Testing	TP	FP	FN	F1
RGBJunNoObs	RGBJunObs32	RGBJunObs	875	1438	11	54.7%
RGBJunNoObs	RGBJunObs32	RGBSeptObs*	1013	1069	554	55.5%

* denotes different weather conditions

Table 5. CS-Flow results on water edge annotated with fixed horizontal line.

Environment adaptation	Tuning	Testing	TP	FP	FN	F1
RGBJunNoObs	RGBJunObs32	RGBJunObs	1433	924	248	71.0%
RGBJunNoObs	RGBJunObs32	RGBSeptObs*	1627	861	685	67.8%

* denotes different weather conditions
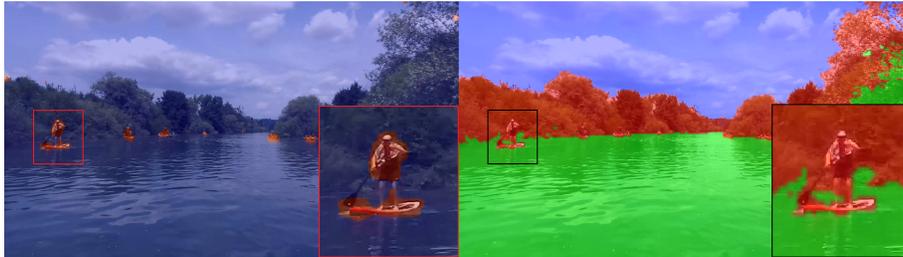
Fig. 3. Comparison of obstacle detection using anomaly detection by PaDiM (left) and WaSR (right) in the same river scene. PaDiM performs better in detecting obstacles not fully surrounded by water, like the paddleman. WaSR's misclassification can cause issues for tracking or motion prediction

6 Discussion and conclusion

To summarize, an important advantage of our cascaded training approach is the ability to split the training into three different phases, where each of them can be served with easily obtainable dataset. In limited geographical domains, common for many USV tasks, we simplify obstacle detection by modeling the obstacle-free environment, redefining obstacle detection as *detection of non-permanent scene items*. Our strategy is viable if *environment adaptation datasets* accurately represent operating conditions. Anomaly detection algorithms require varied training data, but such images can be obtained semi-automatically without labels. Models can be retrained with new data to maintain performance, unlike supervised segmentation models needing precise annotations.

Visual examination shows that performance of some methods might be affected by overshoots in true positive detections (as seen in Figure 2). Further development in anomaly detection methods and more detailed anomaly maps (requiring more powerful hardware) can address this issue, localizing inconsistencies more precisely.

Different nature of detected obstacles. Our framework uses anomaly detection for obstacle detection, which is of a *different nature* than discriminative methods like WaSR [6]. Riverbank features aren't considered obstacles; USVs

Table 6. WaSR results for both types of water edge annotation.

Testing	Fixed line water edge				Polygon water edge			
	TP	FP	FN	F1	TP	FP	FN	F1
RGBJunObs	1679	179	2	94.9%	884	0	2	99.9%
RGBSeptObs	2185	352	127	90.1%	1433	1381	134	65.4%

need accurate GPS/DGPS and updated maps to avoid them. This approach, however, offers segmentation between riverbank and obstacle (unlike WaSR, as in Figure 3), crucial for further analysis. A failure mode occurs when visually similar obstacles appear in unusual places (tree in the middle of the river), and this has to be handled by other sensors (e.g. inexpensive, one beam LIDAR).

References

1. Akcay, S., Atapour-Abarghouei, A., Breckon, T.P.: GANomaly: Semi–supervised Anomaly Detection via Adversarial Training. In: Asian Computer Vision Conference (ACCV). pp. 622–637 (2019)
2. Bergman, L., Hoshen, Y.: Classification-Based Anomaly Detection for General Data. In: International Conference on Learning Representations (ICLR) (2020)
3. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9584–9592 (2019)
4. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: Uninformed Students: Student–Teacher Anomaly Detection With Discriminative Latent Embeddings. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4183–4192 (2020)
5. Bovcon, B., Kristan, M.: A Water-Obstacle Separation and Refinement Network for Unmanned Surface Vehicles. In: IEEE International Conference on Robotics and Automation (ICRA) (2020)
6. Bovcon, B., Kristan, M.: WaSR–A Water Segmentation and Refinement Maritime Obstacle Detection Network. IEEE Transactions on Cybernetics pp. 1–14 (2021)
7. Bovcon, B., Muhovic, J., Perš, J., Kristan, M.: The MaSTr1325 Dataset for Training Deep USV Obstacle Detection Models. In: IEEE International Conference on Intelligent Robots and Systems (IROS). pp. 3431–3438 (2019)
8. Bovcon, B., Muhovič, J., Vranac, D., Mozetič, D., Perš, J., Kristan, M.: MODS–A USV-Oriented Object Detection and Obstacle Segmentation Benchmark. IEEE Transactions on Intelligent Transportation Systems pp. 1–16 (2021)
9. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: NuScenes: A Multimodal Dataset for Autonomous Driving. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11621–11631 (2020)
10. Cane, T., Ferryman, J.: Evaluating Deep Semantic Segmentation Networks for Object Detection in Maritime Surveillance. In: IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). pp. 1–6 (2018)
11. Cohen, N., Hoshen, Y.: Sub-Image Anomaly Detection with Deep Pyramid Correspondences. arXiv preprint arXiv:2005.02357 (2020)
12. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The Cityscapes Dataset for Semantic Urban Scene Understanding. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3213–3223 (2016)
13. Defard, T., Setkov, A., Loesch, A., Audigier, R.: PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization. In: International Conference on Pattern Recognition (ICPR). vol. 12664, pp. 475–489 (2020)
14. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A Large-Scale Hierarchical Image Database. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 248–255 (2009)

15. Fei, Y., Huang, C., Jinkun, C., Li, M., Zhang, Y., Lu, C.: Attribute Restoration Framework for Anomaly Detection. IEEE Transactions on Multimedia pp. 1–12 (2021)

16. Geiger, A., Lenz, P., Urtasun, R.: Are We Ready for Autonomous Driving? The Kitti Vision Benchmark Suite. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3354–3361 (2012)

17. González, A., Fang, Z., Socarras, Y., Serrat, J., Vázquez, D., Xu, J., López, A.M.: Pedestrian Detection at Day/Night Time With Visible and FIR Cameras: A Comparison. Sensors **16**(6), 820 (2016)

18. Gudovskiy, D., Ishizaka, S., Kozuka, K.: CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows. In: IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1819–1828 (2022)

19. Haselmann, M., Gruber, D.P., Tabatabai, P.: Anomaly Detection Using Deep Learning Based Image Completion. In: International Conference on Machine Learning and Applications (ICMLA) (2018)

20. Heidarsson, H.K., Sukhatme, G.S.: Obstacle Detection and Avoidance for an Autonomous Surface Vehicle Using a Profiling Sonar. In: IEEE International Conference on Robotics and Automation (ICRA). pp. 731–736 (2011)

21. Hendrycks, D., Mazeika, M., Kadavath, S., Song, D.: Using Self-Supervised Learning Can Improve Model Robustness and Uncertainty. In: Advances in Neural Information Processing Systems (NIPS) (2019)

22. Hermann, D., Galeazzi, R., Andersen, J., Blanke, M.: Smart Sensor Based Obstacle Detection for High-Speed Unmanned Surface Vehicle. In: IFAC Conference on Manoeuvring and Control of Marine Craft (MCMC). vol. 48, pp. 190–197 (2015)

23. Jeong, M., Li, A.Q.: Efficient LiDAR-based In-water Obstacle Detection and Segmentation by Autonomous Surface Vehicles in Aquatic Environments. In: IEEE International Conference on Intelligent Robots and Systems (IROS). pp. 5387–5394 (2021)

24. Jia, X., Zhu, C., Li, M., Tang, W., Zhou, W.: LLVIP: A Visible-infrared Paired Dataset for Low-light Vision. In: IEEE International Conference on Computer Vision (ICCV). pp. 3496–3504 (2021)

25. Kiefer, B., et al.: 1st Workshop on Maritime Computer Vision (MaCVi) 2023: Challenge Results. In: IEEE Winter Conference on Applications of Computer Vision Workshops (WACVW). pp. 265–302 (2023)

26. Kim, H., Koo, J., Kim, D., Park, B., Jo, Y., Myung, H., Lee, D.: Vision-Based Real-Time Obstacle Segmentation Algorithm for Autonomous Surface Vehicle. IEEE Access **7**, 179420–179428 (2019)

27. Kristan, M., Kenk, V.S., Kovačič, S., Perš, J.: Fast Image-Based Obstacle Detection From Unmanned Surface Vehicles. IEEE Transactions on Cybernetics **46**(3), 641–654 (2016)

28. Lee, S.J., Roh, M.I., Lee, H.W., Ha, J.S., Woo, I.G.: Image-Based Ship Detection and Classification for Unmanned Surface Vehicle Using Real-Time Object Detection Neural Networks. In: International Ocean and Polar Engineering Conference (ISOPE) (2018)

29. Ma, L., Xie, W., Huang, H.: Convolutional Neural Network Based Obstacle Detection for Unmanned Surface Vehicle. Mathematical Biosciences and Engineering **17**, 845–861 (2020)

30. Mai, K.T., Davies, T., Griffin, L.D.: Brittle Features May Help Anomaly Detection. In: Women in Computer Vision workshop at CVPR (2021)

31. Moosbauer, S., Konig, D., Jakel, J., Teutsch, M.: A Benchmark for Deep Learning Based Object Detection in Maritime Environments. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2019)

32. Muhovič, J., Mandeljc, R., Bovcon, B., Kristan, M., Perš, J.: Obstacle Tracking for Unmanned Surface Vessels Using 3-D Point Cloud. IEEE Journal of Oceanic Engineering (2019)

33. Nirgudkar, S., Robinette, P.: Beyond Visible Light: Usage of Long Wave Infrared for Object Detection in Maritime Environment. In: International Conference on Advanced Robotics (ICAR). pp. 1093–1100 (2021)

34. Noroozi, M., Favaro, P.: Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles. In: European Conference on Computer Vision (ECCV). pp. 69–84 (2016)

35. Onunka, C., Bright, G.: Autonomous Marine Craft Navigation: On the Study of Radar Obstacle Detection. In: International Conference on Control Automation Robotics & Vision (ICARCV). pp. 567–572 (2010)

36. Perš, J., Muhovič, J., Bobek, U., Brinšek, Cvenkel, T., Gregorin, D., Mitja, K., Lukek, M., Sedej, N., Kristan, M.: Modular Multi-Sensor System for Unmanned Surface Vehicles. In: International Electrotechnical and Computer Science Conference (ERK) (2021)

37. Qiao, D., Liu, G., Li, W., Lyu, T., Zhang, J.: Automated Full Scene Parsing for Marine ASVs Using Monocular Vision. Journal of Intelligent & Robotic Systems **104**(2) (2022)

38. Rippel, O., Mertens, P., Merhof, D.: Modeling the Distribution of Normal Data in Pre–trained Deep Features for Anomaly Detection. In: International Conference on Pattern Recognition (ICPR) (2020)

39. Rudolph, M., Wehrbein, T., Rosenhahn, B., Wandt, B.: Fully Convolutional Cross-Scale-Flows for Image-based Defect Detection. In: IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1829–1838 (2022)

40. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S.A., Binder, A., Müller, E., Kloft, M.: Deep One-Class Classification. In: International Conference on Machine Learning (ICML). pp. 4393–4402 (2018)

41. Ruiz, A.R.J., Granja, F.S.: A Short-Range Ship Navigation System Based on Ladar Imaging and Target Tracking for Improved Safety and Efficiency. IEEE Transactions on Intelligent Transportation Systems **10**(1), 186–197 (2009)

42. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: f–AnoGAN: Fast Unsupervised Anomaly Detection with Generative Adversarial Networks. Medical Image Analysis **54**, 30–44 (2019)

43. Steccanella, L., Bloisi, D., Castellini, A., Farinelli, A.: Waterline and Obstacle Detection in Images From Low-Cost Autonomous Boats for Environmental Monitoring. Robotics and Autonomous Systems **124**, 103346 (2020)

44. Teledyne: Teledyne FLIR thermal dataset for algorithm training, https://www.flir.eu/oem/adas/adas-dataset-form/

45. Tsai, C.C., Wu, T.H., Lai, S.H.: Multi-Scale Patch-Based Representation Learning for Image Anomaly Detection and Segmentation. In: IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 3992–4000 (2022)

46. Žust, L., Kristan, M.: Learning Maritime Obstacle Detection From Weak Annotations by Scaffolding. In: IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 955–964 (2022)

47. Wang, G., Han, S., Ding, E., Huang, D.: Student-Teacher Feature Pyramid Matching for Unsupervised Anomaly Detection. In: British Machine Vision Conference (BMVC) (2021)

48. Wang, M., Wang, Q., Hong, D., Roy, S.K., Chanussot, J.: Learning Tensor Low-Rank Representation for Hyperspectral Anomaly Detection. IEEE Transactions on Cybernetics **53**(1), 679–691 (2022)
49. Wang, W., Gheneti, B., Mateos, L.A., Duarte, F., Ratti, C., Rus, D.: Roboat: An Autonomous Surface Vehicle for Urban Waterways. In: IEEE International Conference on Intelligent Robots and Systems (IROS). pp. 6340–6347 (2019)
50. Waymo: An autonomous driving dataset (2019)
51. Yang, J., Li, Y., Zhang, Q., Ren, Y.: Surface Vehicle Detection and Tracking with Deep Learning and Appearance Feature. In: International Conference on Control, Automation and Robotics (ICCAR). pp. 276–280 (2019)
52. Yao, L., Kanoulas, D., Ji, Z., Liu, Y.: ShorelineNet: An Efficient Deep Learning Approach for Shoreline Semantic Segmentation for Unmanned Surface Vehicles. In: IEEE International Conference on Intelligent Robots and Systems (IROS). pp. 5403–5409 (2021)
53. Yu, J., Zheng, Y., Wang, X., Li, W., Wu, Y., Zhao, R., Wu, L.: FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows. arXiv preprint arXiv:2111.07677 (2021)
54. Zavrtanik, V., Kristan, M., Skočaj, D.: Reconstruction by Inpainting for Visual Anomaly Detection. Pattern Recognition **112**, 1–9 (2021)
55. Zhan, W., Xiao, C., Wen, Y., Zhou, C., Yuan, H., Xiu, S., Zhang, Y., Zou, X., Liu, X., Li, Q.: Autonomous Visual Perception for Unmanned Surface Vehicle Navigation in an Unknown Environment. Sensors **19**(10) (2019)